Université de Montréal

Routage adaptatif et qualité de service dans les réseaux optiques à commutation de rafales

par

Abdeltouab Belbekkouche

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Thèse présentée à la Faculté des arts et des sciences
en vue de l'obtention du grade de Doctorat
en Informatique

Août, 2010

Université de Montréal
Faculté des arts et des sciences

Cette thèse intitulée :

Routage adaptatif et qualité de service dans les réseaux optiques à commutation de rafales

Présentée par :
Abdeltouab Belbekkouche

a été évaluée par un jury composé des personnes suivantes :

Bernard Gendron, président-rapporteur, représentant du doyen
de la faculté des arts et des sciences
Abdelhakim Hafid, directeur de recherche
Michel Gendreau, co-directeur
Brigitte Jaumard, membre du jury
Martin Maier, examinateur externe

Résumé

Les réseaux optiques à commutation de rafales (OBS) sont des candidats pour jouer un rôle important dans le cadre des réseaux optiques de nouvelle génération. Dans cette thèse, nous nous intéressons au routage adaptatif et au provisionnement de la qualité de service dans ce type de réseaux.

Dans une première partie de la thèse, nous nous intéressons à la capacité du routage multi-chemins et du routage alternatif (par déflection) à améliorer les performances des réseaux OBS, pro-activement pour le premier et ré-activement pour le second. Dans ce contexte, nous proposons une approche basée sur l'apprentissage par renforcement où des agents placés dans tous les nœuds du réseau coopèrent pour apprendre, continuellement, les chemins du routage et les chemins alternatifs optimaux selon l'état actuel du réseau. Les résultats numériques montrent que cette approche améliore les performances des réseaux OBS comparativement aux solutions proposées dans la littérature.

Dans la deuxième partie de cette thèse, nous nous intéressons au provisionnement absolu de la qualité de service où les performances *pire-cas* des classes de trafic de priorité élevée sont garanties quantitativement. Plus spécifiquement, notre objectif est de garantir la transmission sans pertes des rafales de priorité élevée à l'intérieur du réseau OBS tout en préservant le multiplexage statistique et l'utilisation efficace des ressources qui caractérisent les réseaux OBS. Aussi, nous considérons l'amélioration des performances du trafic *best effort*. Ainsi, nous proposons deux approches : une approche basée sur les nœuds et une approche basée sur les chemins. Dans l'approche basée sur les nœuds, un ensemble de longueurs d'onde est assigné à chaque nœud du bord du réseau OBS pour qu'il puisse envoyer son trafic garanti. Cette assignation prend en considération les distances physiques entre les nœuds du bord. En outre, nous proposons un algorithme de sélection des longueurs d'onde pour améliorer les performances des rafales *best effort*. Dans l'approche basée sur les chemins, le provisionnement absolu de la qualité de service est fourni au niveau des chemins entre les nœuds du bord du réseau OBS. À cette fin, nous proposons une approche de routage et d’assignation des longueurs d’onde qui a pour but la réduction du nombre requis de longueurs d’onde pour établir des chemins sans contentions. Néanmoins, si cet objectif ne peut pas être atteint à cause du nombre limité de longueurs d’onde, nous proposons de synchroniser les chemins en conflit sans le besoin pour des équipements additionnels. Là aussi, nous proposons un algorithme de sélection des longueurs d’onde pour les rafales *best effort*. Les résultats numériques montrent que l’approche basée sur les nœuds et l’approche basée sur les chemins fournissent le provisionnement absolu de la qualité de service pour le trafic garanti et améliorent les performances du trafic *best effort*. En outre, quand le nombre de longueurs d’ondes est suffisant, l’approche basée sur les chemins peut accommoder plus de trafic garanti et améliorer les performances du trafic *best effort* par rapport à l’approche basée sur les nœuds.

Mots-clés : Réseau optique à commutation de rafales; Routage; Assignation des longueurs d’onde; Sélection des longueurs d’onde; Apprentissage par renforcement; Optimisation combinatoire; Recherche avec tabou.

Abstract

Optical Burst Switching (OBS) networks are candidates to play an important role in the context of next generation optical networks. In this thesis, we are interested in adaptive routing and quality of service provisioning for these networks.

In the first part of the thesis, we study the capability of multi-path routing and alternative routing (deflection routing) to improve the performance of the OBS network proactively for the former and reactively for the latter. In this context, we propose a reinforcement learning-based approach where learning agents, placed in each OBS node, cooperate to learn, continuously, optimal routing paths and alternative paths according to the current state of the network. Numerical results show that the proposed approach improves the performance of the OBS network compared to existing solutions in the literature.

In the second part of the thesis, we consider the problem of absolute quality of service provisioning for OBS networks where *worst-case* performance of high priority traffic is guaranteed quantitatively. Particularly, we are interested in the loss-free transmission, inside the OBS network, of high priority bursts, while preserving statistical multiplexing gain and high resources utilization of the OBS network. Also, we aim to improve the performance of best effort traffic. Hence, we propose two approaches: (a) the node-based approach; and (b) the path-based approach. In the node-based approach, we propose to assign a set of wavelengths to each OBS edge node that it can use to send its guaranteed traffic. This assignment takes into consideration physical distances between edge nodes. Furthermore, we propose a wavelength selection algorithm to improve the performance of best effort bursts. In the path-based approach, absolute quality of service provisioning is offered at end-to-end path level. To do this, we propose a routing and wavelength assignment approach which aims to reduce the number of wavelengths required to establish contention free paths. Nevertheless, if this objective cannot be reached because of the limited number of wavelengths in each fiber link, we propose an approach to synchronize overlapping paths without the need for additional equipments for synchronization. Here again, we propose a wavelength selection algorithm for best effort bursts. Numerical results show that both the node-based and the path-based approaches successfully provide absolute quality of service provisioning for guaranteed traffic and improve the performance of best effort traffic. Also, path-based approach could accommodate more guaranteed traffic and improve the performance of best effort traffic compared to node-based approach when the number of wavelengths is sufficient.

Keywords : Optical Burst Switching (OBS); Routing; Wavelength assignment; Wavelength Selection; Reinforcement learning; Combinatorial optimization; Tabu search.

Table des matières

Résumé i
Abstract iii
Table des matières v
Liste des tableaux ix
Liste des figures x
Liste des algorithmes xiii
Liste des sigles et abréviations xiv
Remerciements xviii
Chapitre 1 : Introduction 1
  1.1. Contexte et motivations du projet de recherche 1
  1.2. Contributions et organisation de la thèse 7
  1.3. Publications de la thèse 11
Chapitre 2 : Revue de la littérature 14
  2.1. Les réseaux optiques à commutation de rafales 14
    2.1.1. Architecture et fonctionnement 14
    2.1.2. L'assemblage et le désassemblage des rafales 18
    2.1.3. Signalisation 20
    2.1.4. L'ordonnancement des rafales 21
    2.1.5. Résolution des contentions 23
  2.2. Le routage multi-chemins 25
  2.3. Le routage alternatif 31
  2.4. La sélection des longueurs d'onde 35
  2.5. Le provisionnement absolu de la qualité de service 38
Chapitre 3 : Novel reinforcement learning-based approaches to reduce loss probability in buffer-less OBS networks 46
  3.1. Introduction 47
  3.2. Related work 49

3.3. Reinforcement learning 52
3.4. The proposed deflection routing scheme 53
  3.4.1. The proposed scheme 54
  3.4.2. Exploration, convergence and overhead analysis 58
  3.4.3. Insufficient offset time problem 59
3.5. The proposed multi-path routing approach 60
  3.5.1. The routing scheme 60
  3.5.2. Exploration, convergence and overhead analysis 62
  3.5.3. Loopless routing, out-of-order delivery and insufficient offset time 63
3.6. The integrated approach 66
3.7. Simulation results and analysis 67
3.8. Conclusion and future work 74
Chapitre 4 : Topology-aware wavelength partitioning for DWDM OBS networks: A novel approach for absolute QoS provisioning 76
4.1. Introduction 77
4.2. Related work 79
  4.2.1. Absolute QoS provisioning 79
  4.2.2. Wavelength assignment 81
4.3. Optimization-based topology-aware wavelength partitioning approach 82
  4.3.1. OTWP description 83
  4.3.2. Exact formulation 85
  4.3.3. Complexity analysis 88
  4.3.4. Tabu search algorithm 89
4.4. Absolute fair QoS differentiation scheme 93
  4.4.1. The operation of AFQD 93
  4.4.2. Best effort traffic wavelength assignment scheme 96
  4.4.3. Starvation problem for BE traffic 97
4.5. Wavelength borrowing protocol 98
4.6. Numerical results 102

4.6.1. Results of CPLEX and tabu search algorithm 103
4.6.2. Results of AFQD 104
4.6.3. Results of BETWA 110
4.6.4. Results of WBP 113
4.7. Concluding remarks 115
Chapitre 5 : Path-based QoS Provisioning for Optical Burst Switching Networks 117
5.1. Introduction 118
5.2. Related work 121
5.2.1. QoS provisioning 121
5.2.2. Synchronization in OBS networks 122
5.3. Path-based QoS provisioning 123
5.3.1. Overview 124
5.3.2. Asynchronous PQP 124
5.3.3. Synchronous PQP 125
5.3.4. Path-based best effort wavelength selection 126
5.4. Routing and wavelength assignment approach 127
5.4.1. Routing paths optimization 127
5.4.2. Wavelength assignment 131
5.5. Path-based synchronous transmission scheme 136
5.5.1. Overview 136
5.5.2. Capacity constraint problem 137
5.5.3. The synchronization protocol 142
5.5.4. Synchronization issues 146
5.6. Numerical results 146
5.6.1. Results of ILP models (A, B, C and D) and tabu search algorithm 148
5.6.2. Results of asynchronous PQP 151
5.6.3. Results of synchronous PQP 153
5.6.4. Results of PBEWS 156
5.7. Concluding remarks 159

Chapitre 6 : Conclusion et travaux futurs 160
  6.1. Contributions et résultats de la thèse... 160
  6.2. Perspectives et travaux futurs 163
Bibliographie... 165

Liste des tableaux

Tableau 1. Comparaison des paradigmes de la commutation optique. 5

Tableau 2. Récapitulatif des travaux sur le routage multi-chemins dans les réseaux OBS. 31

Tableau 3. Récapitulatif des travaux sur le routage alternatif dans les réseaux OBS. 35

Tableau 4. Récapitulatif des travaux sur le provisionnement absolu de la QdS dans les réseaux OBS. 45

Table 1. Comparison of Tabu Search (TS) and CPLEX results. 104

Table 2. The number of required wavelengths to establish non-overlapping paths using shortest path routing for twelve of SNDlib instances. 120

Table 3. Results of routing paths optimization ILP model (A). 149

Table 4. Results of wavelength assignment ILP model (B) and Tabu Search (TS) algorithm 150

Table 5. Results of ILP model C and ILP model D for Janos-us-ca instance when varying the number of wavelengths 150

Liste des figures

Figure 1. Évolution des réseaux optiques selon Battestilli *et al.* [3] 2
Figure 2. Organigramme de l'organisation de la thèse. 11
Figure 3. Architecture d'un réseau OBS. 14
Figure 4. Les fonctionnalités principales au long du chemin d'une rafale. 17
Figure 5. Architecture et stratégies d'assemblage des rafales selon Gauger *et al.* [20]....... 18
Figure 6. Signalisation dans les réseaux OBS.. 21
Figure 7. Routage multi-chemins vs. routage du plus court chemin.................................. 26
Figure 8. Variantes du déploiement des solutions du routage multi-chemins. 27
Figure 9. An Example that shows a part of node *x*'s deflection table................................. 56
Figure 10. The operation of RLDRS at an OBS core node.. 57
Figure 11. An example that shows a part of node *x*'s Q-table. 61
Figure 12. The operation of RLMR at an OBS core node. 62
Figure 13. The operation of IRLRCR at an OBS core node. 67
Figure 14. NSFNET topology with 14 nodes (a) and regular torus topology with 4 x 4 nodes (b)... 68
Figure 15. Loss probability vs. load on NSFNET with 64 wavelengths............................ 70
Figure 16. Mean burst end-to-end delay (number of hops) vs. load on NSFNET with 64 wavelengths... 71
Figure 17. Loss probability vs. load on regular 4 x 4 topology with 64 wavelengths. 72
Figure 18. Mean burst end-to-end delay (number of hops) vs. load on regular 4 x 4 topology with 64 wavelengths. 73
Figure 19. Loss probability vs. number of wavelengths on NSFNET. 74
Figure 20. Mean burst end-to-end delay (number of hops) vs. number of wavelengths on NSFNET.. 74
Figure 21. An example of wavelength partitioning using OTWP scheme. 84
Figure 22. 14-nodes NSFNET topology. 88
Figure 23. The operation of AFQD to schedule a LS burst at source node. 95

Figure 24. The order in which BETWA searches an available wavelength at node *i* with wavelength interval *i'* 97
Figure 25. The operation of AFQD to schedule a BE burst at source node........ 97
Figure 26. The fields of WBP packets 101
Figure 27. WBP interactions........ 101
Figure 28. (a) 4 x 4 regular topology and (b) 15-nodes ring topology. 103
Figure 29. Loss probability vs. load for AFQD and LAUC-VF on NSFNET with 64 wavelengths........ 106
Figure 30. BE traffic loss probability vs. load on NSFNET with 64wavelengths........ 107
Figure 31. LS traffic blocking probability on NSFNET with 64wavelengths........ 107
Figure 32. Loss probability of AFDQ and LAUC-VF on NSFNET when varying the number of wavelengths. 108
Figure 33. Blocking probability of LS and loss probability of BE and LAUC-VF on 4 x 4 regular topology with 64 wavelengths. 109
Figure 34. Blocking probability of LS and loss probability of BE and LAUC-VF on 15-nodes ring topology with 64 wavelengths........ 110
Figure 35. BETWA: loss probability vs. load with NSFNET and 64 wavelengths........ 111
Figure 36. BETWA: loss probability vs. number of wavelengths with NSFNET topology when load is fixed to 50%........ 112
Figure 37. BETWA: loss probability vs. number of wavelengths with 4 x 4 nodes regular topology and 64 wavelengths........ 112
Figure 38. BETWA: loss probability vs. number of wavelengths with ring topology and 64 wavelengths........ 113
Figure 39. Blocking probability of LS with and without WBP on NSFNET and 64 wavelengths........ 114
Figure 40. Blocking probability of LS with and without WBP on regular topology with 64 wavelengths........ 114
Figure 41. Blocking probability of LS with and without WBP on ring topology and 64 wavelengths........ 115

Figure 42. Operation of PQP (synchronous and asynchronous operation modes). 125
Figure 43. The order of searching an available wavelength for a BE burst on path p_i assigned a wavelength w_j. 127
Figure 44. Example of four paths sharing the fiber link (A, B).. 128
Figure 45. Example of three connected components of paths assigned wavelength 2. 137
Figure 46. Example of a connected component conflict sub-graph of paths operating on the same wavelength. 139
Figure 47. Example of a synchronization tree. 144
Figure 48. The main fields of Sync packet. 145
Figure 49. NSFNET topology. 147
Figure 50. Overall loss probability vs. load for PQP and AFQD. 152
Figure 51. Blocking probability for LS traffic. 152
Figure 52. Comparison of PQP and TSA 154
Figure 53. QoS differentiation of synchronous PQP. 154
Figure 54. Blocking probability of LS traffic with synchronous PQP.............................. 155
Figure 55. Impact of varying timeslot duration on loss (a) and delay (b)......................... 156
Figure 56. Impact of varying buffer size on loss (a) and delay (b). 157
Figure 57. Loss probability of PBEWS vs. BETWA with 64 wavelengths...................... 158
Figure 58. Loss probability of PBEWS vs. BETWA when varying the number of wavelengths. 158

Liste des algorithmes

Algorithm 1. Loopless forwarding algorithm. 65
Algorithm 2. The Tabu Search (TS) algorithm. 91
Algorithm 3. The Construction Method (CM) heuristic. 92
Algorithm 4. Tabu search algorithm for wavelength assignment. 134
Algorithm 5. Initialization greedy heuristic for wavelength assignment. 135

Liste des sigles et abréviations

ACK	Acknowledgment
AFQD	Absolute Fair Quality of service Differentiation scheme
ATM	Asynchronous Transfer Mode
BETWA	Best Effort Traffic Wavelength Assignment scheme
CLDR	Contention-based Limited Deflection Routing
DT	Deflection Table
DTMC	discrete-time Markov chain
DTT	Distance To Threshold
DVLP	Dynamic Virtual Lambda Partitioning
FDL	Fiber Delay Line
IP	Internet Protocol
IRLRCR	Integrated Reinforcement Learning-based Routing and Contention Resolution
JET	Just Enough Time
JIT	Just In Time
LAUC	Last Available Unused Channel
LBDR	Load Balancing Deflection Routing algorithm
LLN	Least Loaded Node
LUW	Least Used Wavelength
NACK	Negative Acknowledgment
NSFNET	National Science Foundation NETwork
OADM	Optical Add Drop Multiplexer
OBS	Optical Burst Switching
OBSRWA	OBS Routing and Wavelength Assignment
OCS	Optical Circuit Switching
OEO	Opto-Électro-Optique
OPS	Optical Packet Switching
OT	Offset Time

OTSI	Optical TimeSlot Interchangers
OTWP	Optimization-based Topology-aware Wavelength Partitioning approach
OXC	Optical cross-Connect
PBEWS	Path-based Best Effort Wavelength Selection scheme
PQP	Path-based QoS Provisioning
PRDR	Prioritized Random Deflection Routing
PST	Path-based Synchronous Transmission scheme
PWA	Priority-based Wavelength Assignment
QdS	Qualité de Service
QoS	Quality of Service
QWS	Q-learning algorithm for Wavelength Selection
RAM	Random Access Memory
RLDRS	Reinforcement Learning-based Deflection Routing Scheme
RLMR	Reinforcement Learning-based Multi-path Routing
SI	Suitability Index
SOBS	Synchronous Optical Burst Switching
SP	Shortest Path
SPDR	Shortest Path Deflection Routing
SPF	Shortest Path First
TCP	Transmission Control Protocol
TE	Traffic Engineering
TPS	Tree-based Path Synchronization
TSA	Tree-based Slot Allocation
VCR	Virtual Channel Reservation
VF	Void Filling
WAN	Wide Area Network
WBP	Wavelength Borrowing Protocol
WDM	Wavelength Division Multiplexing

Glossaire des traductions

Acknowledgment	Acquittement
Bit error rate	Taux d'erreurs sur les bits
Bit rate	Débit binaire
Burst	Rafale
Feed-back	Retour
Fiber Delay Line	Ligne de retardement
Guard band	Bande de garde
In-band	Intra-bande
Leaky Bucket	Seau percé
Lightpath	Chemin lumineux
Multi-path	Multi-chemins
Offset Time	Temps de décalage
Insufficient Offset Time	Temps de décalage insuffisant
Optical Burst Switching	Commutation de rafales optiques
Optical Circuit Switching	Commutation de circuits optiques
Optical Packet Switching	Commutation de paquets optiques
Out of band	Hors bande
Overhead	Surplus
Wavelength Division Multiplexing	Multiplexage en longueur d’onde
Random Access Memory	Mémoire vive
Round Trip Time	Temps de rotation
Out-of-Order delivery	Livraison en dehors de l'ordre
Jitter	Gigue
Starvation	Famine
Store-and-forward	Stockage-et-retransmission
Timeslot	Intervalle de temps

À ma famille.

Remerciements

Je remercie tous ceux et celles qui ont participé de prés ou de loin à l'accomplissement de cette thèse.

Je remercie mon directeur de recherche Abdelhakim Hafid, professeur à l'université de Montréal, pour la qualité de son encadrement et l'intérêt qu'il a démontré à l'égard de cette thèse. Je le remercie pour vos conseils et directives pertinentes qui ont été de grande importance pour la réalisation de ce projet de recherche.

Mes remerciements s'adressent à Michel Gendreau, professeur à l'école polytechnique de Montréal, pour les discussions fructueuses et les précieux conseils qu'il a toujours donné au cours de cette thèse.

Je tiens à remercier Monsieur Bernard Gendron, professeur à l'université de Montréal, Madame Brigitte Jaumard, professeure à l'université Concordia et Monsieur Martin Maier, professeur à l'Institut National de la Recherche scientifique, d'avoir accepté de faire partie du jury pour examiner les travaux de recherche de cette thèse.

Aussi, je tiens à remercier les collègues au Laboratoire des Réseaux de Communications (LRC) et au Centre Interuniversitaire de Recherche sur les Réseaux d'Entreprise, la Logistique et le Transport (CIRRELT) pour leur amitié et pour la collaboration fructueuse que nous avons pu établir.

Enfin, je remercie ma femme, mes parents et mes frères qui m'ont donné l'énergie, l'ambition et le soutient pour en arriver jusque là.

Chapitre 1 : Introduction

Nous commençons ce chapitre par présenter le contexte et les motivations de la thèse. Ensuite, nous décrivons ses contributions et son organisation. Enfin, nous présentons les publications (articles de revue et de conférence) publiés ou soumis au cours de cette thèse.

1.1. Contexte et motivations du projet de recherche

Avec l'émergence de nouvelles applications et de nouveaux services (tels que les applications multimédias, la voix-sur-IP, la vidéo-sur-demande, le calcul de grille, etc.), la demande des utilisateurs d'Internet sur la bande passante est devenue de plus en plus croissante. Cette situation nécessite le recours à la fibre optique comme support de transmission de données. Cette dernière offre un potentiel énorme de bande passante ainsi que des caractéristiques hors pair tels qu'un faible taux d'atténuation et de distorsion du signal et un faible taux d'erreur sur les bits [1]. Ce potentiel en bande passante de la fibre optique est mieux exploité en faisant recours à la technologie du multiplexage en longueur d'onde (*Wavelength Division Multiplexing* [WDM]). Cette technologie consiste à subdiviser (virtuellement) la fibre optique en plusieurs canaux orthogonaux, appelés longueurs d'onde, capables d'envoyer les signaux optiques indépendamment les uns des autres et sans interférences. Cependant, si cette solution résout le problème de transmission des données, la commutation des données dans les nœuds du cœur du réseau demeure, quant-à-elle, problématique. Ceci est notamment vrai si les données doivent être transformées dans le domaine électronique à l'entrée de chaque nœud du cœur et retransformées encore dans le domaine optique à sa sortie. Il s'agit de la transformation Opto-Électro-Optique (OEO). Un réseau optique dans lequel chaque nœud du cœur effectue cette transformation pour le trafic en transit est dit réseau *opaque*. Ainsi, à cause de la grande capacité de transmission du domaine optique par rapport à la capacité de traitement du domaine électronique, les nœuds du cœur vont former des goulots d'étranglement dans le réseau optique. Aussi, la consommation de l'énergie et la diffusion de la chaleur sont deux facteurs négatifs dans le fonctionnement des réseaux opaques. C'est dans ce contexte que des dispositifs de commutation optique tels que les *Optical Add Drop Multiplexers* (OADMs) et les *Optical Cross-Connects* (OXCs) ont vu le jour pour donner naissance aux réseaux *tout-optiques* (*transparents*) [2]. Dans les réseaux tout-optiques, le trafic reste dans le domaine optique de sa source à sa destination sans subir des transformations OEO dans les nœuds du cœur du réseau. Ceci est possible grâce aux dispositifs de commutation optique.

La Figure 1 illustre l'évolution historique des réseaux optiques et leur passage des réseaux opaques (où la partie optique se limite à des liens WDM point à point) aux réseaux tout-optiques avec les trois paradigmes de commutation suivants :

- Commutation de circuits optiques (*Optical Circuit Switching* [OCS]);
- Commutation de rafales optiques (*Optical Burst Switching* [OBS]);
- Commutation de paquets optiques (*Optical Packet Switching* [OPS]).

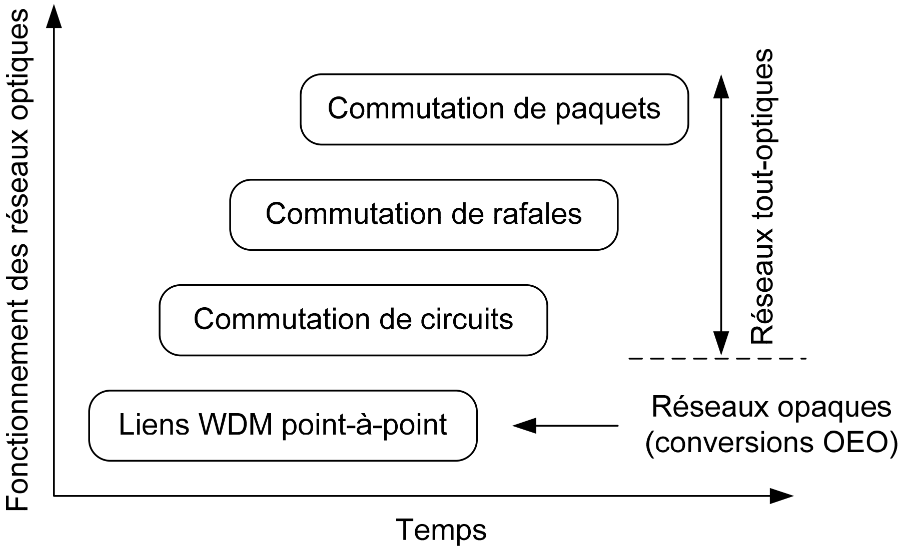


Figure 1. Évolution des réseaux optiques selon Battestilli *et al.* [3]

La commutation de circuits, aussi connue sous le nom de routage de longueur d'onde, consiste à établir une connexion optique, appelée *lightpath* [4], sur une longueur d'onde pour envoyer le trafic entre un nœud source et un nœud destination. La durée de cette connexion peut varier de l'ordre des millisecondes à l'ordre des jours, des semaines et même des mois. À noter que la majorité des réseaux tout-optiques déployés aujourd'hui sont de type commutation de circuits. Bien que la commutation de circuits garantie un niveau de qualité de service élevé pour le trafic accepté dans le réseau et ne nécessite pas des commutateurs optiques ultra rapides, cette technique souffre de deux problèmes majeurs. Le premier problème est celui du temps de latence pour établir une connexion optique entre une source et une destination. En effet, l'envoi du trafic sur une connexion optique ne peut commencer avant que son nœud source n'envoie une requête d'établissement de la connexion et ne reçoive un acquittement positif qui confirme l'établissement de la connexion et la réservation des ressources nécessaires. Ce temps d'aller-retour de la requête et de la réponse peut s'avérer très grand, notamment, si le diamètre du réseau est assez grand. Le second problème est celui de l'utilisation des ressources. En effet, dans les réseaux optiques à commutation de circuits, les ressources utilisées par une connexion (y compris la longueur d'onde sur chaque lien) ne peuvent pas être partagées avec d'autres connexions. Sachant que la bande passante requise par une seule application (ou un seul utilisateur) est souvent bien au dessous de la bande passante disponible dans une longueur d'onde, ceci produira une sous utilisation des ressources et une probabilité de blocage élevée pour les requêtes des nouvelles connexions.

Le paradigme de la commutation de paquets optiques est similaire à celui des réseaux électroniques à commutation de paquets (tels que les réseaux IP) [5]. Ainsi, un paquet optique, composé d'un entête contenant les informations de contrôle et d'une charge utile représentant les données transportées par ce paquet, est commuté dans chaque nœud intermédiaire dans le domaine optique. Pour cela, l'entête du paquet doit être traité pour effectuer les opérations de routage et de réservation des ressources. Le traitement de l'entête peut être effectué soit (a) dans le domaine optique; soit (b) dans le domaine électronique après son extraction de son paquet et sa transformation dans le domaine électronique. Dans ce dernier cas, il faudra retarder la charge utile du paquet (dans le domaine optique) en utilisant des tampons optiques. À noter que le retardement de la charge utile d'un paquet optique, dans chaque nœud intermédiaire, pour traiter son entête augmentera son délai de bout en bout. Malheureusement, l'avancement technologique actuel dans le domaine des équipements optiques ne permet pas la réalisation de la solution (a) et ne permet pas la réalisation efficace de la solution (b) [2]. De plus, la petite taille des paquets optiques nécessite des commutateurs optiques ultra-rapides (de l'ordre des nanosecondes, ou même des picosecondes), ce qui complique d'avantage la réalisation des réseaux optiques à commutation de paquets. Ainsi, il est clair que l'implantation des réseaux optiques à commutation de paquets[1] n'est pas envisageable dans le futur proche.

La commutation de rafales optiques, qui est le sujet d'intérêt dans cette thèse, est considérée comme étant un compromis entre la commutation de circuits et la commutation de paquets. En effet, la commutation de rafales rassemble leurs avantages et évite leurs inconvénients. Dans la commutation de rafales optiques, l'unité commutable à l'intérieur du réseau est un super-paquet, appelé *rafale*, qui est composé de plusieurs paquets (IP, ATM, etc.). La rafale est assemblée dans le nœud du bord source, désassemblée dans le nœud du bord destination et commutée dans les nœuds intermédiaires dans le domaine optique. Un paquet de contrôle est envoyé, avant l'envoi de la rafale, d'une durée appelée temps de décalage (*offset time*). Le rôle du paquet de contrôle est la réservation des ressources nécessaires à sa rafale dans les nœuds intermédiaires. Il est envoyé sur une longueur d'onde dédiée aux paquets de contrôle et il subit des transformations OEO dans chaque nœud intermédiaire. Cette séparation entre le plan de données et le plan de contrôle est l'une des points forts des réseaux OBS (et d'autres types de réseaux tels que les réseaux optiques à commutation de paquets RINGO et HORNET [2]) puisqu'elle permet de bénéficier, à la fois, de la transparence du plan de données et de la capacité du traitement électronique du plan de contrôle. La granularité intermédiaire de la commutation de rafales (entre le circuit et le paquet) réduit le besoin pour des commutateurs ultra-rapides qui sont nécessaires pour la commutation de paquets. Aussi, la commutation de rafales possède le

---

[1] Notons ici qu'il existe une technique prometteuse pour implémenter les réseaux optiques à commutation de paquets qui s'appelle *Optical Label Switching* (OLS) [2] et qui est similaire aux réseaux optiques à commutation de rafales.

multiplexage statistique et le niveau élevé d’utilisation des ressources de la commutation de paquets, ainsi que la transparence (du moins au niveau du plan de données) et la simplicité de réalisation de la commutation de circuits. Les réseaux OBS seront décrits plus en détail dans le chapitre 2 de ce manuscrit.

Le Tableau 1 présente une comparaison des trois paradigmes de commutation optique (commutation de circuits, de paquets et de rafales) en termes de l'utilisation de la bande passante, la latence de l'établissement des connexions, la rapidité requise des commutateurs optiques et la complexité de réalisation. Nous constatons que la commutation de rafales possède une utilisation élevée de la bande passante et une faible latence pour transmettre les données. En outre, la commutation de rafales se situe entre la commutation de circuits et la commutation de paquets en termes de la rapidité requise des commutateurs et en termes de la complexité de réalisation avec la technologie actuelle des équipements optiques.

Tableau 1. Comparaison des paradigmes de la commutation optique.

Paradigme de commutation	**Utilisation de la bande passante**	**Latence des connexions**	**Rapidité des commutateurs**	**Complexité de réalisation**
Circuit	Faible	Élevée	Faible	Faible
Paquet	Élevée	Sans connexions	Élevée	Élevée
Rafale	Élevée	Faible	Moyenne	Moyenne

Les caractéristiques attractives des réseaux OBS, en prenant en considération l’avancement technologique actuel et dans le future proche, font de ces réseaux des candidats pour jouer un rôle important dans le cadre des réseaux optiques de nouvelle génération. En effet, les réseaux OBS peuvent jouer un rôle dans les différents niveaux de la hiérarchie des réseaux optiques tels que les réseaux optique étendus [2] ou encore le cœur des réseaux optiques métropolitains [6, 7]. Cependant, les réseaux OBS souffrent de certains défis et problèmes tel que le problème des contentions de longueur d'onde. Une contention de longueur d'onde apparaît quand les réservations de deux ou plusieurs rafales qui doivent emprunter un même port de sortie sur la même longueur d'onde se chevauchent. Ce problème est directement lié au manque d'une mémoire optique efficace analogue à la mémoire vive (*Random Acces Memory* [RAM]) dans les réseaux électroniques. D'autres défis sont liés à l'absence de standards pour les réseaux OBS et l'absence de mécanismes efficaces de provisionnement de la qualité de service analogues à ceux utilisés dans les réseaux électroniques à commutation de paquets. En effet, les mécanismes de provisionnement de la qualité de service des réseaux électroniques (tel que les réseaux IP) sont, essentiellement, basés sur le principe du *stockage-et-retransmission* (*store-and-forward)* qui se base sur les mémoires tampons électroniques efficaces qui n'ont pas d'analogues dans le domaine optique.

Ainsi, pour que les réseaux OBS soient acceptés comme une solution viable pour les réseaux optiques de nouvelle génération, il faudra d'abord trouver des solutions efficaces à ces problèmes. C'est ce que nous essayerons de faire dans cette thèse en identifiant deux problématiques majeures dans les réseaux OBS. La première problématique est celle de l'amélioration des performances des réseaux OBS en l'absence des tampons optiques et des dispositifs de conversion de longueur d'onde qui sont, actuellement, jugés onéreux, inefficaces ou technologiquement immatures. La deuxième problématique est celle du provisionnement absolu de la qualité de service pour le trafic de priorité élevée, toujours en l'absence des tampons optiques et des dispositifs de conversion de longueur d'onde. Plus spécifiquement, nous nous intéressons à garantir la transmission sans pertes aux rafales de priorité élevée à l'intérieur du réseau OBS, quelque soit la topologie de ce dernier. Notons que le réseau OBS à l'étude, dans cette thèse, est caractérisé par son aspect économique et sa simplicité de réalisation avec une technologie des équipements optiques bien maîtrisée actuellement.

1.2. Contributions et organisation de la thèse

Comme indiqué dans la section précédente, nous nous penchons dans cette thèse sur deux problématiques majeures dans les réseaux OBS, à savoir, l'amélioration des performances de ces réseaux en utilisant le routage adaptatif ainsi que le provisionnement absolu de la qualité de service en l'absence des tampons optiques et des dispositifs de conversion de longueur d'onde. Dans ce cadre, nous présentons trois contributions dont la première concerne le routage adaptatif et les deux autres concernent le provisionnement absolu de la qualité de service.

Pour la première contribution, nous explorons la capacité du routage adaptatif et distribué à améliorer les performances du réseau OBS. Pour cela, nous proposons un mécanisme proactif pour réduire le nombre de contentions en utilisant le routage multi-chemins. Nous appelons ce mécanisme *Reinforcement Learning Based Multi-path Routing* (RLMR). Aussi, nous proposons un mécanisme réactif adaptatif pour résoudre les contentions en utilisant le routage alternatif (par déflection). Nous appelons ce mécanisme *Reinforcement Learning Based Deflection Routing Scheme* (RLDRS). RLMR et RLDRS sont basés sur une approche d'apprentissage par renforcement où des agents placés dans chaque nœud du réseau OBS apprennent, continuellement, le chemin optimal pour router une rafale de son nœud source à son nœud destination et le chemin alternatif optimal pour router une rafale d'un nœud intermédiaire où elle a été impliquée dans une contention à son nœud destination. La combinaison de RLMR et RLDRS donne lieu à une approche intégrée de routage et de résolution des contentions, que nous appelons *Integrated Reinforcement Learning-based Routing and Contention Resolution* (IRLRCR). IRLRCR améliore de façon substantielle les performances du réseau OBS par rapport aux travaux antérieurs trouvés dans la littérature. Cette contribution fera l'objet du chapitre 3 de cette thèse. Le contenu de ce chapitre a été publié dans la revue *Computer Networks :* A. Belbekkouche, A. Hafid and M. Gendreau. *Novel reinforcement learning-based approaches to reduce loss probability in buffer-less OBS networks.* Computer Networks, Vol. 53(12): p. 2091-2105. 2009 [8].

La deuxième contribution est consacrée au problème du provisionnement absolu de la qualité de service. Plus spécifiquement, nous nous intéressons à la capacité du réseau OBS à garantir une transmission sans pertes à l'intérieur du réseau OBS pour le trafic sensible aux pertes, et ce, quelque soit la topologie du réseau. Pour cela, nous proposons une approche, appelée *Absolute Fair Quality of service Differentiation scheme (AFQD)*. AFQD assigne un ensemble de longueurs d'onde (une ou plusieurs) à chaque nœud du bord du réseau OBS en prenant en considération la topologie du réseau, c.-à-d., les distances physiques entre les nœuds du bord. Sachant que les contentions, dans le cœur du réseau, ne peuvent apparaître entre des rafales provenant du même nœud source, chaque nœud pourra transmettre son trafic garanti sans pertes à l'intérieur du réseau OBS. Ainsi, nous proposons une approche d'assignation des longueurs d'onde aux nœuds que nous appelons *Optimization-based Topology-aware Wavelength Partitioning approach (OTWP)*. OTWP modélise le problème d'assignation sous forme d'un modèle d'optimisation linéaire et utilise un algorithme de recherche avec tabou pour résoudre les instances de grande taille efficacement. Afin d'éviter de perdre les caractéristiques attractives des réseaux OBS tels que le multiplexage statistique et le niveau élevé d'utilisation des ressources, les rafales non garanties, de type *best effort*, peuvent utiliser n'importe quelle longueur d'onde. De plus, pour réduire la probabilité de perte du trafic *best effort*, nous proposons un algorithme de sélection des longueurs d'onde, appelé *Best Effort Traffic Wavelength Assignment scheme (BETWA)*. BETWA se base sur l'assignation des longueurs d'onde aux nœuds, effectuée par OTWP, pour sélectionner une longueur d'onde à une rafale *best effort* dans son nœud source. BETWA vise à maximiser l'isolation des trafics *best effort* des différents nœuds du réseau OBS. Pour rendre AFQD adaptatif aux différents modèles de trafic, nous proposons un protocole d'adaptation de l'assignation des longueurs d'onde, appelé *Wavelength Borrowing Protocol (WBP)*. Dans WBP, les nœuds du bord du réseau OBS peuvent échanger des longueurs d'onde qui leur ont été assignées pour s'adapter aux variations du trafic. Les résultats numériques montrent que AFQD est capable de garantir le provisionnement absolu de la qualité de service et de réduire, efficacement, la probabilité de perte du trafic *best effort* tout en s'adaptant aux variations du trafic dans le réseau OBS.

Cette contribution fera l’objet du chapitre 4 de cette thèse. Le contenu de ce chapitre a été accepté pour publication dans la revue *Computer Networks :* A. Belbekkouche, A. Hafid, M. Tagmouti and M. Gendreau. *Topology-aware wavelength partitioning for DWDM OBS networks: A novel approach for absolute QoS provisioning*. To appear in Computer Networks [9].

La troisième contribution est consacrée au problème du provisionnement absolu de la qualité de service, et plus spécifiquement, la garantie de la transmission sans pertes à l’intérieur du réseau OBS pour le trafic garanti. Cependant, pour cette contribution, nous explorons la possibilité de donner cette garantie au niveau d’un chemin de bout en bout au lieu d’un nœud. Nous appelons cette approche *Path-based QoS Provisioning* (PQP). Nous considérons, d’abord, le problème de réduire le nombre requis de longueurs d’onde pour établir des chemins non-chevauchants entre chaque paire de nœuds du bord du réseau OBS. Dans une telle configuration, les cas potentiels de contention entre les chemins (où deux chemins partagent la même longueur d’onde sur le même lien) sont absents et, par conséquent, les rafales envoyées sur ces chemins ont la garantie d’atteindre leurs destinations sans pertes. Ainsi, nous proposons une approche, appelée *OBS Routing and Wavelength Assignment* (OBSRWA), qui utilise un modèle d’optimisation linéaire pour trouver la configuration des chemins qui minimise le nombre requis de longueurs d’onde et un algorithme de recherche avec tabou pour assigner les longueurs d’onde aux chemins. L’objectif de l’algorithme de recherche avec tabou est de minimiser les conflits entre les chemins (les cas potentiels de contention) ou, idéalement, les éliminer complètement. Néanmoins, si ce dernier cas n’est pas possible, les conflits sont répartis de façon équitable sur l’ensemble des longueurs d’onde. Dans ce cas, nous proposons de synchroniser la transmission des chemins en conflit afin de permettre une transmission sans pertes au trafic garanti à l’intérieur du réseau OBS. À cette fin, nous proposons une approche de synchronisation, appelée *Path-based Synchronous Transmission scheme* (PST). PST se base sur un protocole de synchronisation des chemins en conflit, appelé *Tree-based Path Synchronization* (TPS). Contrairement aux travaux trouvés dans la littérature qui prônent le paradigme OBS synchronisé [10-16], PST ne requiert pas d'équipements additionnels pour effectuer la synchronisation. Aussi, pour maximiser la quantité du trafic garanti que chaque chemin peut transporter, PST utilise deux modèles d'optimisation linéaires pour déterminer la capacité maximale de tous les chemins à synchroniser et pour allouer à chaque chemin ses intervalles de temps pour envoyer du trafic garanti. Comme pour la deuxième contribution, nous proposons, là aussi, un algorithme de sélection des longueurs d'onde pour le trafic *best effort*, appelé *Path-based Best Effort Wavelength Selection scheme* (PBEWS), qui a pour but de conserver le multiplexage statistique et le niveau élevé de l'utilisation des ressources dans les réseaux OBS et de réduire la probabilité de perte du trafic *best effort*. Les résultats numériques montrent que PQP peut accommoder plus de trafic garanti que AFQD [9] sur la topologie NSFNET avec un nombre suffisant de longueurs d'onde (13 longueurs d'onde dans ce cas). De plus, comparativement à BETWA (utilisé par AFQD), la probabilité de perte du trafic *best effort* est réduite par PBEWS. Cette contribution fera l'objet du chapitre 5 de cette thèse. Le contenu de ce chapitre a été soumis pour publication à la revue *Journal of Lightwave Technology :* A. Belbekkouche, A. Hafid, M. Gendreau and M. Tagmouti. *Path-based QoS Provisioning for Optical Burst Switching Networks*. Submitted to IEEE/ACM Transactions on Networking [17].

La thèse est organisée comme suit. Après ce chapitre introductif, le chapitre 2 présente une brève description des réseaux OBS et une revue de la littérature sur les sujets abordés dans cette thèse, à savoir, le routage multi-chemins, le routage alternatif, l'assignation des longueurs d'onde et le provisionnement de la qualité de service dans les réseaux OBS. Ensuite, le chapitre 3 présente la première contribution sur le routage adaptatif dans les réseaux OBS. Les chapitres 4 et 5 présentent, respectivement, la deuxième et la troisième contribution. Ils ont pour objet le provisionnement de la qualité de service dans les réseaux OBS. Finalement, le chapitre 6 trace les conclusions de cette thèse et identifie quelques pistes de recherche pour les travaux futurs.

La Figure 2 illustre un organigramme des contributions et de l'organisation de cette thèse.

1.3. Publications de la thèse

La liste des articles de revues et des articles de conférences rédigés au cours de cette thèse est la suivante :

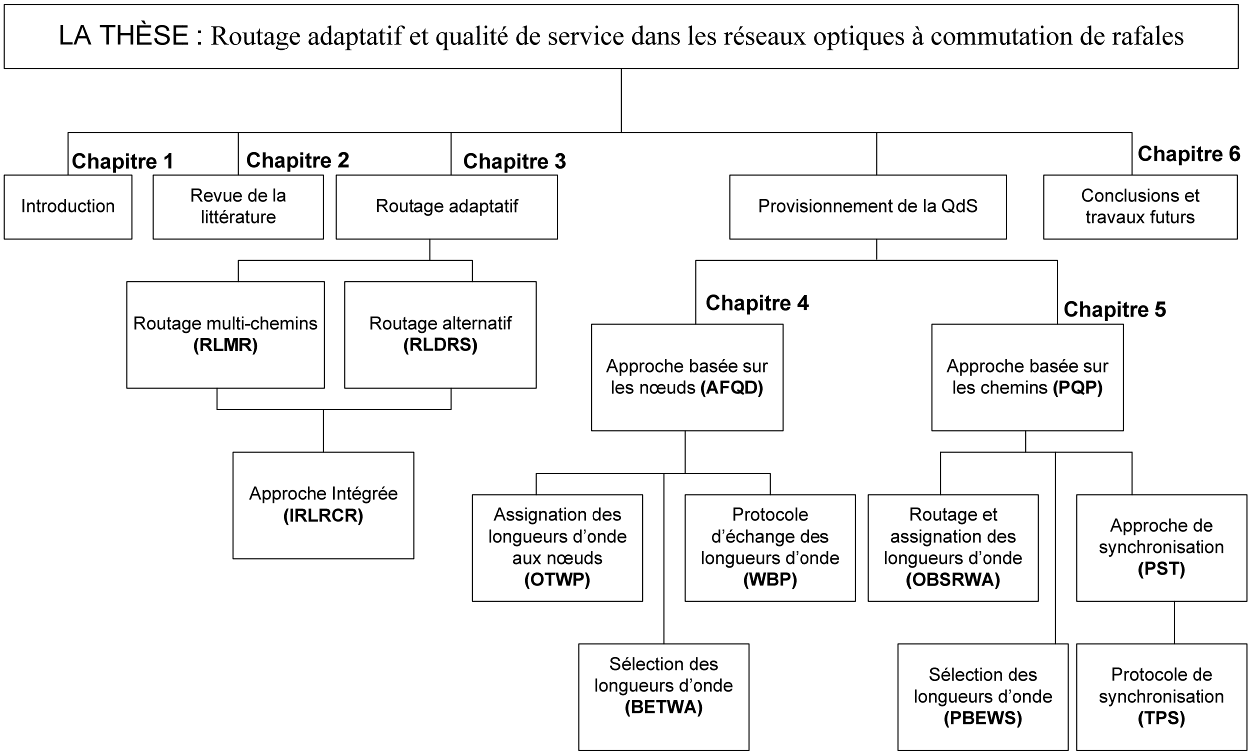


Figure 2. Organigramme de l'organisation de la thèse.

1. A. Belbekkouche, A. Hafid, and M. Gendreau, M. Tagmouti. *Path-based QoS for Optical Burst Switching Networks*. Submitted to IEEE/ACM Transactions on Networking.

2. A. Belbekkouche, J. Rezgui and A. Hafid. *Wireless Mesh and Optical Burst Switching Convergence for a Novel Metropolitan Area Network Architecture*. To appear in Computer Networks (accepted August 2010).

3. A. Belbekkouche, A. Hafid, M. Tagmouti and M. Gendreau. *Topology-Aware Wavelength Partitioning For DWDM OBS Networks: A Novel Approach for Absolute QoS Provisioning*. To appear in Computer Networks (accepted June 2010).

4. A. Belbekkouche, A. Hafid, M. Gendreau. *Novel reinforcement learning-based approaches to reduce loss probability in buffer-less OBS networks*. Computer Networks, Vol. 53(12), pages: 2091-2105, 2009.

5. A. Belbekkouche, A. Hafid, M. Gendreau and M. Tagmouti. *RWA and Synchronization to Provide Absolute QoS for OBS Networks*. in Proceedings of the 10th INFORMS Telecommunications Conference (extended abstract), Montréal, Canada, 2010.

6. A. Belbekkouche, A. Hafid, M. Tagmouti and M. Gendreau. *A Novel Formulation for Routing and Wavelength Assignment Problem in OBS Networks*. in Proceedings of IEEE International Conference on Communications (ICC'10), Cape Town, South Africa, 2010.

7. A. Belbekkouche, J. Rezgui and A. Hafid. *QoS provisioning for Wireless Mesh and Optical Burst Switching Convergence*. in Proceedings of IEEE Wireless Communications and Networking Conference (WCNC'10), Sydney, Australia, 2010.

8. A. Belbekkouche, A. Hafid, M. Tagmouti and M. Gendreau. *An Absolute and Fair QoS Differentiation Scheme for DWDM OBS Networks*. in Proceedings of IEEE Global Telecommunications Conference (Globecom'09), Honolulu, USA , 2009.

9. A. Belbekkouche, A. Hafid and M. Gendreau. *Adaptive Routing and Contention Resolution Approaches for Buffer-less OBS Networks*. in Proceedings of ICST International Conference on Broadband Communications, Networks, and Systems (Broadnets'09), 2009.

10. A. Belbekkouche, A. Hafid and M. Gendreau. *A Reinforcement Learning-based Deflection Routing Scheme for Buffer-less OBS Networks*. in Proceedings of IEEE Global Telecommunications Conference (Globecom'08), New Orleans, USA, 2008.

11. A. Belbekkouche and A. Hafid. *An Adaptive Reinforcement Learning-based Approach to Reduce Blocking Probability in Buffer-less OBS Networks*. in Proceedings of IEEE International Conference on Communications (ICC'07), Glasgow, UK, 2007.

Chapitre 2 : Revue de la littérature

2.1. Les réseaux optiques à commutation de rafales

Dans cette section, nous présentons un aperçu sur les réseaux OBS. Dans ce cadre, nous présentons l'architecture *typique* et les éléments qui composent un réseau OBS. Par la suite, nous présentons les fonctionnalités principales dans un réseau OBS, à savoir, le processus d'assemblage des rafales dans les nœuds du bord, la signalisation tout au long du chemin emprunté par les rafales, la réservation des ressources pour les rafales et la résolution des contentions dans les nœuds du cœur du réseau OBS. D'autres aperçus sur les réseaux OBS peuvent être trouvés dans [2, 18-21].

2.1.1. Architecture et fonctionnement

Une rafale est un *super-paquet* qui encapsule plusieurs paquets provenant des réseaux clients adjacents au réseau OBS. La taille de la rafale peut varier de quelques Kilo Octets à quelques Méga Octets (la taille d'un fichier, par exemple). La Figure 3 illustre l'architecture d'un réseau OBS dans lequel nous distinguons les éléments suivants :

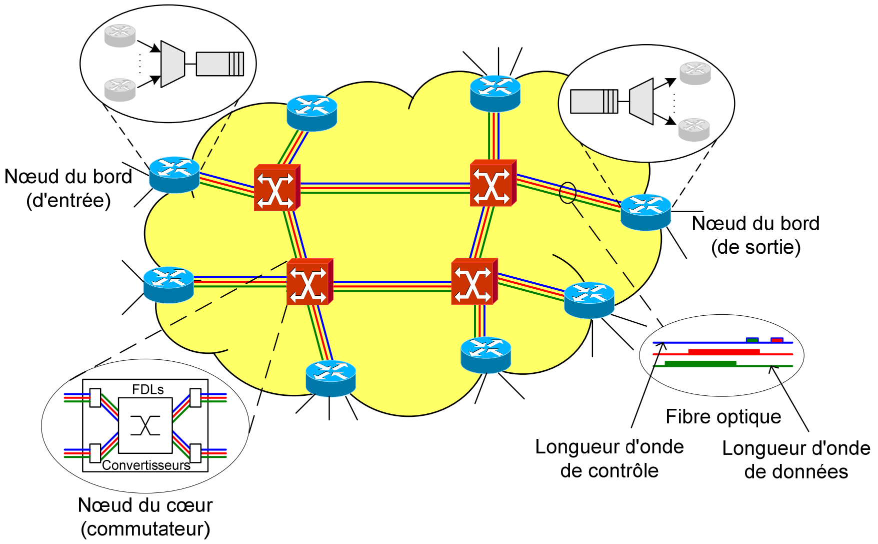


Figure 3. Architecture d'un réseau OBS.

- *Le nœud du bord du réseau :* dans ce nœud, les rafales sont assemblées (avec les paquets des réseaux clients) et stockées (si nécessaire) dans le domaine électronique. Pour chaque rafale assemblée, un paquet de contrôle est généré dans ce nœud et envoyé vers la destination de la rafale afin d'effectuer les réservations nécessaires des ressources tout au long du chemin qui sera emprunté par la rafale. Généralement, un temps de décalage (*Offset Time*) sépare le paquet de contrôle de sa rafale pour compenser le délai apporté par les transformations OEO et le traitement électronique du paquet de contrôle dans chaque nœud intermédiaire. Aussi, ce nœud sert à désassembler les rafales dont il est la destination et à faire passer les paquets qui les composent aux réseaux clients (des routeurs IP ou ATM par exemple).

- *Les liens optiques :* il s'agit des liens en fibre optique qui relient tous les nœuds du réseau entre eux. Il peut y avoir une ou plusieurs fibres optiques qui composent un seul lien entre deux nœuds pour offrir la caractéristique bidirectionnelle au lien, ou encore, pour augmenter la capacité de celui-ci. Chaque fibre optique est subdivisée (virtuellement) en plusieurs canaux orthogonaux appelés longueurs d'onde. Ceci peut être réalisé grâce à la technologie du multiplexage en longueurs d'onde (*Wavelength Division Multiplexing* [WDM]). Dépendamment des variantes technologiques, une longueur d'onde peut posséder une capacité en bande passante de quelques gigabits par seconde (Gbps) (jusqu'à une dizaine par exemple), ce qui donne à la fibre entière une capacité dans l'ordre des térabits par seconde (Tbps) si la fibre opère avec une centaine de longueurs d'onde, par exemple. Pour réaliser la séparation du plan de données et du plan de contrôle, une ou plusieurs longueurs d'onde sont dédiées aux paquets de contrôle. Le reste des longueurs d'onde est utilisé pour les rafales de données.

- *Le nœud du cœur du réseau :* c'est un commutateur optique équipé d'une unité de traitement électronique pour traiter les paquets de contrôle. En effet, le paquet de contrôle subit une transformation OEO à chaque nœud du cœur pour extraire les informations relatives à la rafale correspondante (destination, durée, etc.) et effectuer les réservations des ressources nécessaires à celle-ci. Ainsi, le commutateur optique est configuré pour commuter la rafale à son arrivée sur un port d'entrée au port de sortie approprié sans avoir recours à la transformation OEO; d'où la transparence du plan de données du réseau OBS. Les lignes à retardement (*Fibre Delay Lines* [FDLs]), qui servent à retarder les rafales dans le domaine optique, ainsi que les convertisseurs de longueur d'onde, qui servent à assigner à la rafale une longueur d'onde différente de celle avec laquelle elle arrive au nœud, sont des éléments optionnels dans les nœuds du cœur.

- *La ligne à retardement (Fiber Delay Line [FDL]) :* c'est un segment de fibre optique qui joue le rôle d'un tampon optique utilisé pour retarder les rafales pour une durée prédéterminée. Les FDLs peuvent être utilisées comme moyen de résolution des contentions entre les rafales (quand deux ou plusieurs rafales doivent entamer le même port de sortie sur la même longueur d'onde et dont les durées de transmission se chevauchent). Aussi, les FDLs peuvent être utilisées comme une alternative (ou du moins comme un moyen d'ajustement) au temps de décalage qui sépare le paquet de contrôle de sa rafale. Dans ce cas, la rafale est retardée dans chaque nœud intermédiaire en utilisant les FDLs afin de permettre au nœud de transformer le paquet de contrôle dans le domaine électronique et de traiter les informations qu'il transporte. Cependant, la durée de retardement d'une FDL est fixe et dépend de la taille de celle-ci et de la vitesse de la lumière dans la fibre (~200.000 km/s), ce qui rend les FDLs inflexibles. De plus, si la FDL est partagée entre un ensemble de ports d'entrée (ou de sortie), elle peut devenir à son tour sujet au problème de contention entre les rafales provenant de ces ports.

- *Le convertisseur de longueur d'onde :* c'est un dispositif qui sert à assigner à une rafale, qui a une longueur d'onde λ_1 en entrée d'un nœud, une autre longueur d'onde λ_2 en sortie du nœud. La conversion de longueur d'onde est efficace pour la résolution des contentions entre les rafales et rend l'assignation des longueurs d'onde dans le nœud du bord moins critique. Cependant, les convertisseurs de longueur d'onde requièrent une technologie complexe qui n'a pas encore atteint sa maturité. En outre, ces dispositifs de conversion de longueur d'onde sont onéreux, ce qui augmente le coût des réseaux OBS lorsqu'ils sont installés dans tous les nœuds [22].

Du point de vue fonctionnel, les paquets du réseau client sont assemblés dans les nœuds du bord dans des rafales, selon leurs destinations et, éventuellement, selon des contraintes de qualité de service. Avant d'envoyer une rafale, un paquet de contrôle est créé et envoyé pour effectuer la signalisation et les réservations des ressources nécessaires pour la rafale. Le paquet de contrôle contient les informations de sa rafale, tels que sa taille et sa destination. En général, un temps de décalage (OT) sépare l'envoie de la rafale de son paquet de contrôle pour donner le temps nécessaire à ce dernier d'effectuer les réservations tout au long du chemin vers la destination de la rafale. Le paquet de contrôle subit une transformation OEO à chaque nœud intermédiaire entre la source et la destination. La rafale est envoyée après le temps de décalage et elle est commutée dans le domaine optique dans chaque nœud intermédiaire jusqu'à sa destination. La rafale est désassemblée dans son nœud de destination et les paquets qui la composent sont envoyés à leurs réseaux clients de destination. La Figure 4 montre les fonctionnalités principales dans un réseau OBS tout au long du chemin d'une rafale. Nous constatons que l'ordonnancement des rafales et la résolution des contentions sont effectués dans les nœuds intermédiaires du réseau. Dans ce qui suit, nous présenterons ces fonctionnalités avec plus de détails.

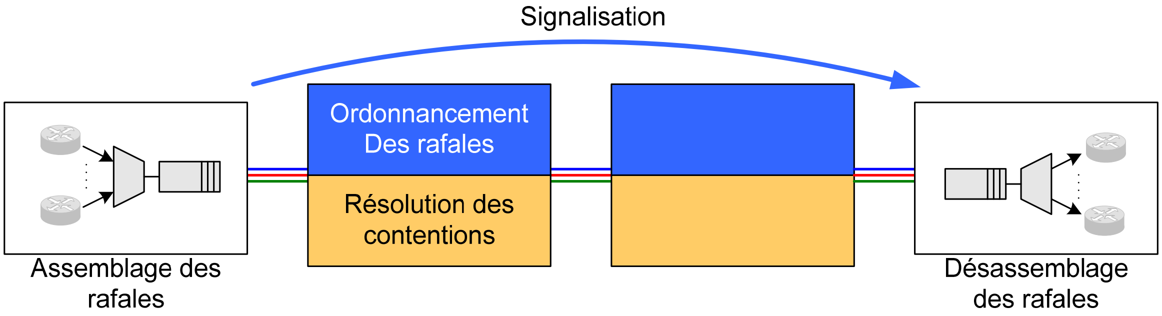


Figure 4. Les fonctionnalités principales au long du chemin d'une rafale.

2.1.2. L’assemblage et le désassemblage des rafales

L’assemblage des rafales s’effectue dans les nœuds du bord du réseau OBS. Les paquets qui entrent dans un nœud du bord sont d’abord classifiés selon leurs destinations et, éventuellement, d’autres critères tels que la qualité de service des paquets. Ces paquets sont assemblés dans des files d’attentes jusqu’à ce que l’algorithme d’assemblage décide d’arrêter le processus d’assemblage et de créer une nouvelle rafale. À ce moment, un paquet de contrôle contenant les informations de la rafale est crée et il est transféré avec sa rafale à un autre tampon pour ordonnancer leurs transmissions dans le réseau OBS. La Figure 5 illustre l’architecture d’un assembleur de rafales et les stratégies d’assemblage des rafales.

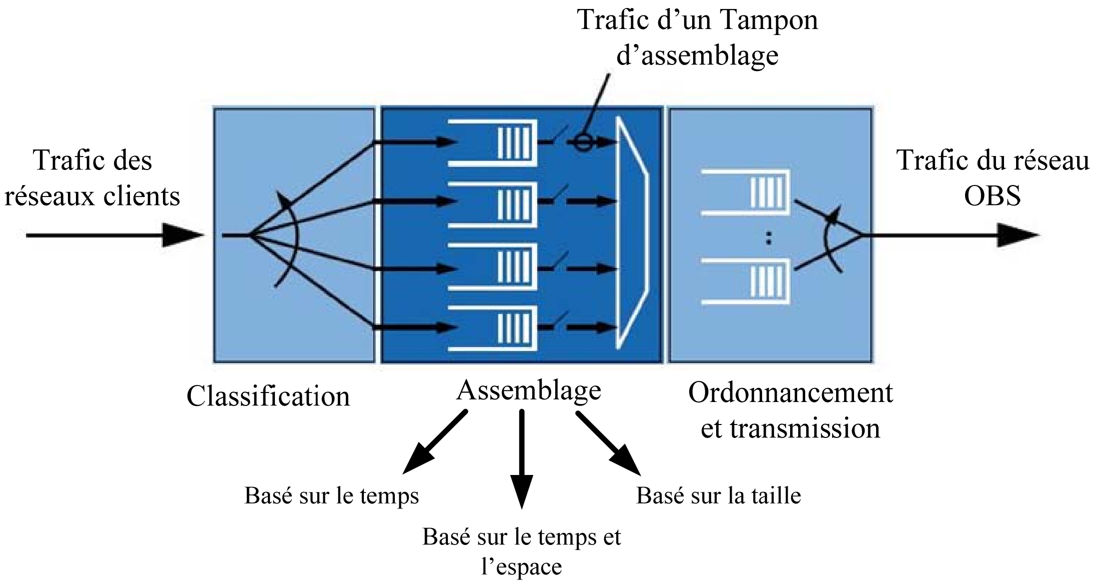


Figure 5. Architecture et stratégies d’assemblage des rafales selon Gauger *et al.* [20].

Les stratégies d’assemblage des rafales peuvent être :

- *basés sur le temps :* ces stratégies utilisent un seuil de temps pour limiter le temps de composition de la rafale, et ainsi, limiter le temps d’attente des paquets dans le nœud du bord du réseau OBS. Ces stratégies sont utiles pour le trafic temps réel qui exige une qualité de service stricte en termes de délais [23].

- *basés sur la taille :* ces stratégies utilisent des seuils pour la taille minimale ou maximale d'une rafale en prenant en considération la quantité de données ou le nombre de paquets dans la rafale. Ces mécanismes sont motivés par des contraintes de réalisation telles que la taille minimale de l'unité commutable dans les nœuds du cœur ou encore la capacité de stockage des tampons d'assemblage [24-26].

- *hybrides :* ces stratégies combinent les paramètres des deux premières stratégies, à savoir, le seuil du temps et le seuil de la taille pour contrôler, en même temps, le délai d'assemblage et la taille des rafales [27, 28].

D'autres stratégies plus avancées adaptent les paramètres du processus d'assemblage à la variation de la charge du trafic ou à l'état du réseau [29, 30].

L'assemblage des rafales est un concept fondamental et attractif des réseaux OBS. En effet, ce processus qui s'effectue à la frontière entre le domaine électronique et le domaine optique possède les avantages suivants [20] :

- ✓ Les paquets sont assemblés dans les rafales qui sont des super-paquets de taille plus large, ce qui réduit le surplus (*overhead*) de la bande de garde (*guard band*) entre les paquets successifs à l'intérieur du réseau.

- ✓ Le débit binaire (*bit rate*) élevé du réseau optique nécessite des commutateurs ultra-rapides dans les nœuds du cœur du réseau optique. Dans les réseaux OBS, l'assemblage des rafales réduit cette nécessité puisque chaque rafale est commutée comme une seule unité à l'intérieur du réseau OBS.

- ✓ L'encapsulation de plusieurs paquets dans une seule rafale renforce la transparence du réseau OBS vis-à-vis ses réseaux clients. En effet, des paquets originaires de plusieurs réseaux clients différents peuvent être assemblés dans une même rafale. Par la suite, ces paquets, dans leur rafale, subiront le même traitement à l'intérieur du réseau OBS.

En outre, les processus d'assemblage et de désassemblage s'effectuent dans le domaine électronique, ce qui réduit leur complexité grâce aux capacités de traitement de ce dernier. Néanmoins, il faut noter que l'assemblage d'une rafale dans le nœud source et son désassemblage dans le nœud destination ajoutent des tâches supplémentaires spécifiques aux réseaux OBS.

Après la transmission des rafales à travers le réseau OBS et dans le nœud du bord destination, la rafale est désassemblée et les paquets qui la composent sont envoyés à leurs réseaux clients de destination.

2.1.3. Signalisation

La transmission d'une rafale à travers le réseau OBS nécessite la réservation d'un ensemble de ressources (telle qu'une longueur d'onde dans chaque lien) et la configuration des commutateurs des nœuds intermédiaires pour commuter la rafale lors de son passage, d'un port d'entrée au port de sortie approprié. Ainsi, une signalisation du passage de la rafale doit être effectuée tout au long du chemin de sa source à sa destination. Dans un réseau OBS, cette signalisation est effectuée hors bande (*out of band*) par les paquets de contrôle qui sont envoyés sur une ou plusieurs longueurs d'onde dédiées aux informations de contrôle.

Dans les réseaux OBS, la signalisation à une seule passe, de type *Tell-And-Go* (TAG), est adoptée [18]. Ainsi, chaque rafale est envoyée après le délai de décalage (qui la sépare de son paquet de contrôle) sans attendre la réception d'un message d'acquittement qui confirme la réservation des ressources pour cette rafale jusqu'à sa destination. L'avantage de la signalisation à une seule passe est qu'elle évite le délai d'attente du message d'acquittement positif qui peut aller jusqu'à la durée de rotation du réseau OBS. De plus, une signalisation à deux passes peut nécessiter une grande capacité de stockage dans les nœuds du bord du réseau OBS puisque les rafales doivent attendre la réception du message d'acquittement dans le domaine électronique.

La Figure 6 schématise le mécanisme de signalisation à une seule passe pour les réseaux OBS. Dans cette figure, les Δ_i représentent le temps de la transformation OEO et du traitement électronique que le paquet de contrôle subit dans chaque nœud intermédiaire.

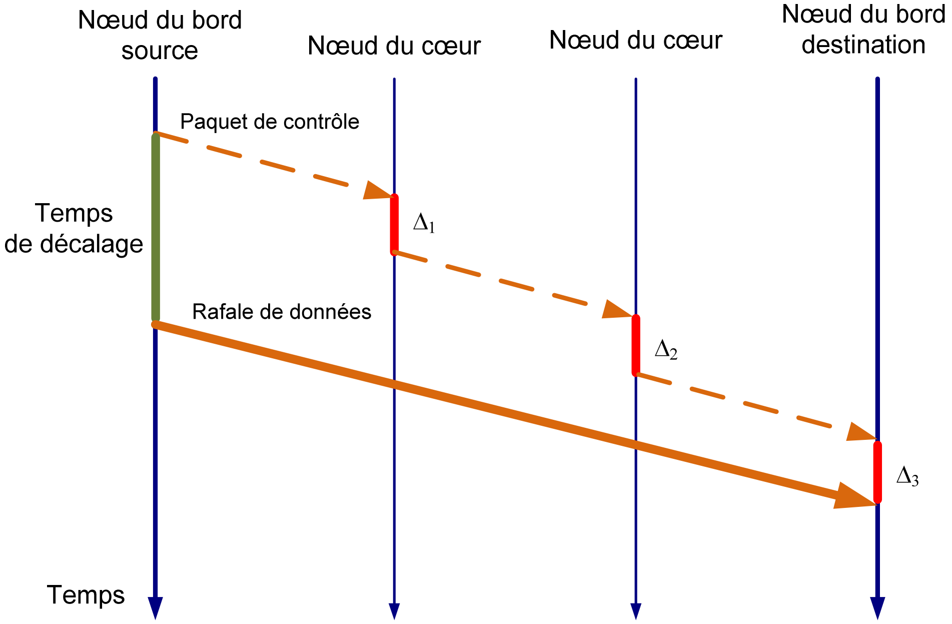


Figure 6. Signalisation dans les réseaux OBS.

2.1.4. L'ordonnancement des rafales

Dans l'ordonnancement des rafales nous nous intéressons à la durée de réservation, à la sélection et à la gestion des ressources au niveau d'un seul nœud OBS.

Pour le début de réservation et la libération des ressources, nous distinguons les quatre variantes suivantes :

- *Réservation immédiate* : la ressource est réservée immédiatement après le traitement du paquet de contrôle. Cette méthode est simple à implanter, cependant, elle entraine un gaspillage des ressources.

- *Réservation retardée* : la réservation de la ressource prend effet au moment de l'arrivée de la rafale (après le délai qui sépare le paquet de contrôle de celle-ci). Cette méthode est complexe à implanter car elle nécessite une estimation du temps exact de l'arrivée de la rafale. Cependant, elle est caractérisée par une bonne utilisation des ressources.

- *Libération implicite* : la ressource est libérée sans un message explicite qui demande sa libération. Idéalement, la libération est effectuée immédiatement après le passage de la rafale en question. Comme pour le cas de la réservation retardée, cette méthode est complexe à implanter mais elle est caractérisée par une bonne utilisation des ressources.

- *Libération explicite* : un message explicite est nécessaire pour libérer la ressource. Comme pour la réservation immédiate, cette méthode est simple à implanter mais elle souffre d'une mauvaise utilisation des ressources.

La combinaison de ces méthodes de réservation et de libération des ressources peuvent donner lieu à des variantes des protocoles de réservation des ressources. Le protocole *Just In Time* (JIT) [31] est à réservation immédiate et à libération explicite, ce qui rend son implantation simple. Cependant, JIT souffre d'une mauvaise utilisation des ressources. D'autre part, le protocole *Just Enough Time* (JET) [18] est à réservation retardée et à libération implicite. Ce qui le rend optimal en termes de l'utilisation des ressources mais complexe en termes de l'implantation.

Dans cette thèse, nous adoptons JET comme protocole de réservation des ressources, puisque celui-ci optimise l'utilisation de la bande passante qui est la ressource essentielle dans les réseaux optiques. D'autre part, la complexité de la réalisation du protocole JET est de plus en plus réduite avec l'avancement technologique dans les ordonnanceurs électroniques des commutateurs optiques.

La sélection des ressources est une phase importante dans le processus d'ordonnancement des rafales. Par exemple, pour assigner une longueur d'onde à la rafale au niveau d'un nœud du bord du réseau OBS, une stratégie de sélection est nécessaire. Les stratégies les plus simples sont : *first-fit* et *round robin*. Avec *first-fit*, la première ressource disponible rencontrée est sélectionnée. Avec *round robin*, les ressources sont sélectionnées une par une, périodiquement et dans un ordre prédéfini. D'autres stratégies plus complexes choisissent la ressource qui est devenue disponibles le plus récemment dans le temps. Ces stratégies sont généralement appelées *Latest Available Unscheduled Channel* (LAUC) [32]. Outre les stratégies principales de sélection des ressources, d'autres stratégies sont proposées dans la littérature. Celles-ci ont pour objectif de réduire le temps d'ordonnancement, d'optimiser le processus de sélection ou de réduire la fragmentation des ressources. Pour ce dernier objectif, une stratégie qui combine LAUC avec le remplissage des vides entre les réservations successives (*Void Filling* [VF]) est proposée et appelée LAUC-VF [32]. Un aperçu des mécanismes d'ordonnancement principaux peut être trouvée dans [33].

2.1.5. Résolution des contentions

La signalisation à une seule passe adoptée pour les réseaux OBS ainsi que le manque de mémoires optiques efficaces sont à l'origine du problème des contentions entre les rafales dans les nœuds du cœur du réseau OBS. Une contention apparaît quand deux ou plusieurs rafales doivent emprunter le même port de sortie (même lien optique) sur la même longueur d'onde et que les durées de réservation de ces rafales se chevauchent [34, 35].

Pour résoudre les contentions après leur apparition, les quatre solutions qui suivent ont été proposées dans la littérature. Sans perte de généralité, nous considérons le cas d'une contention entre deux rafales :

- *Conversion de longueur d'onde* : l'une des deux rafales impliquées dans la contention se voit assigner une longueur d'onde différente de celle qui lui a été assignée avant l'apparition de la contention. Cette solution suppose la présence des convertisseurs de longueur d'onde dans les nœuds du cœur du réseau. Cette solution enlève *la contrainte de la continuité de longueur d'onde* qui stipule que la rafale doit être envoyée sur une seule longueur d'onde tout au long de sa transmission dans le réseau OBS. Comme déjà mentionné dans la section 2.1, outre le coût élevé des convertisseurs de longueur d'onde, ils restent, aujourd'hui, des dispositifs complexes qui n'ont pas encore atteint leur maturité technologique nécessaire pour que leur déploiement dans les réseaux tout-optiques soit répandu. La conversion de longueur d'onde peut être (1) totale : c.-à-d., de n'importe quelle longueur d'onde à n'importe quelle autre longueur d'onde; (2) partielle : c.-à-d., seulement un sous ensemble fixe de longueurs d'onde peut être converti à un autre sous ensemble fixe de longueurs d'onde; (3) à domaine limité : c.-à-d., une longueur d'onde ne peut être convertie qu'à ses longueurs d'onde voisines (une ou plusieurs); et finalement (4) clairsemée : c.-à-d., les convertisseurs de longueur d'onde ne sont placés que dans quelques nœuds du cœur du réseau OBS [2].

- *Lignes à retardement (FDLs)*: l'une des deux rafales impliquées dans la contention est retardée en utilisant une ligne à retardement. Comme nous avons déjà mentionné dans la section 2.1, les lignes à retardement manquent de flexibilité puisqu'elles ne peuvent retarder les rafales que pour des durées fixes. Aussi, une ligne à retardement peut devenir à son tour une ressource sur laquelle les rafales entrent en contentions.

- *Routage alternatif* : l'une des deux rafales impliquées dans la contention est envoyée sur un lien de sortie alternatif où la longueur d'onde de cette rafale est disponible. L'idée du routage alternatif (par déflection) est d'utiliser le réseau entier comme une ressource partagée pour résoudre les contentions. Le routage alternatif est simple à implanter et moins onéreux que les autres solutions (notamment les convertisseurs de longueur d'onde et les FDLs). De plus, le routage alternatif est efficace pour réduire le taux de perte quand la charge du trafic dans le réseau est faible ou modérée. Cependant, le routage par déflection est moins efficace quand la charge du trafic est trop élevée.

- *Segmentation* : alors que les solutions précédentes préservent l'intégrité des rafales, la segmentation consiste à envoyer la partie non concernée par la contention, de l'une des deux rafales impliquées dans la contention, de façon normale et à éliminer la partie concernée par la contention. Cette solution ajoute des fonctionnalités complexes aux nœuds du cœur du réseau OBS pour effectuer et notifier l'opération de segmentation.

Dans cette thèse, nous nous intéressons, particulièrement, au routage alternatif en proposant une approche adaptative à l'état du réseau OBS qui sélectionne le lien de sortie alternatif optimal (en termes du taux de perte et du délai) dans les cas de résolution des contentions (cf. chapitre 3).

2.2. Le routage multi-chemins

Les approches de résolution des contentions décrites dans la section 2.1.5 sont des approches *réactives*, c.-à-d., qu'elles n'interviennent que pour résoudre les contentions après leurs apparitions. Pourtant, il existe des approches, dites *proactives*, qui ont pour objectif de réduire le taux d'apparition des contentions, et par conséquent, réduire le taux de perte des rafales dans les nœuds du cœur du réseau OBS. Les approches proactives sont, généralement, implantées dans les nœuds du bord et servent à envoyer chaque rafale sur le chemin optimal vers sa destination (dans un contexte de routage multi-chemins) et à lui assigner la longueur d'onde optimale [36]. L'optimalité d'un chemin ou d'une longueur d'onde peut être mesurée par le niveau de pertes des rafales, le délai moyen de bout en bout des rafales, une métrique qui combine les deux [37] ou n'importe quelle métrique significative.

Contrairement au routage à chemin unique (*single path routing*) qui consiste à envoyer le trafic entre une paire (source, destination) donnée sur un seul chemin (en général, sur le plus court chemin), le routage multi-chemin (*multi-path routing*) consiste à envoyer le trafic entre une paire (source, destination) sur un ensemble prédéterminé de chemins. Notons ici que le routage du plus court chemin est l'approche de routage la plus utilisée dans les réseaux OBS car elle permet de minimiser les délais et d'optimiser l'utilisation des ressources pour transmettre le trafic à travers le réseau OBS. Par exemple, dans la Figure 7, nous avons deux chemins : C1 et C2 entre les nœuds S et D. Le chemin C1 qui est composé de trois sauts nécessite moins de réservations de ressources que le chemin C2 qui est composé de quatre sauts. Cependant, l'amélioration des performances qui peut être apportée par le routage multi-chemins peut justifier son utilisation dans les réseaux OBS.

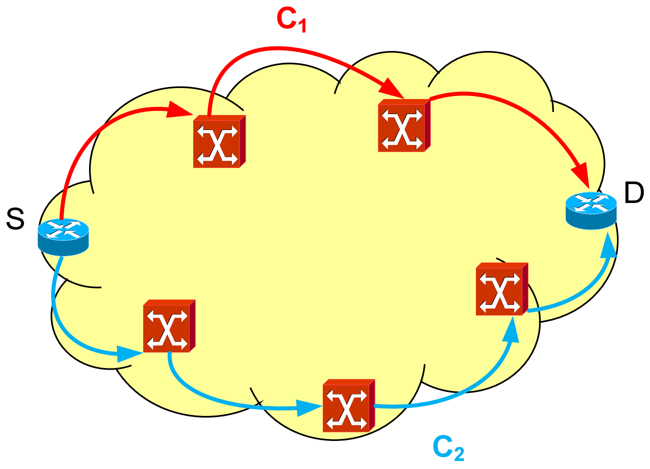


Figure 7. Routage multi-chemins vs. routage du plus court chemin.

Les différents travaux sur le routage multi-chemins dans les réseaux OBS peuvent être classifiés selon la façon de calculer les chemins, le déploiement de la solution (l'emplacement où les décisions sur le routage sont prises), ainsi que la manière de sélectionner les chemins. Le calcul des chemins peut être *statique* ou *adaptatif* à l'état du réseau (topologie ou modèle du trafic). Le déploiement de la solution du routage peut être effectué dans une entité centrale (déploiement centralisé), dans certains nœuds du réseau (en particulier dans les nœuds source, dans ce cas le déploiement est quasi-centralisé) ou encore dans tous les nœuds du réseau (déploiement distribué). La manière de sélectionner les chemins peut être fixe, probabiliste ou basée sur le classement adaptatif des chemins ou des liens. La Figure 8 illustre les variantes du routage multi-chemins du point de vue du déploiement de la solution du routage.

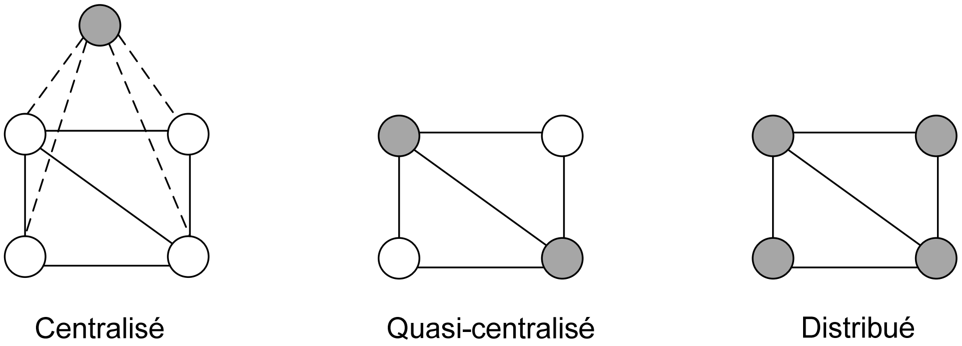


Figure 8. Variantes du déploiement des solutions du routage multi-chemins.

Yang *et al.*, dans [38], proposent un ensemble de stratégies de sélection des chemins dans le contexte du routage multi-chemins dans les réseaux OBS. Pour cela, chaque nœud du bord maintient une liste fixe de chemins vers chaque autre nœud du bord. Cette liste est triée selon les informations les plus récentes sur le niveau de congestion dans le réseau et la stratégie à utiliser. Une stratégie est dite *pure* ou *hybride* dépendamment de la façon avec laquelle elle sélectionne le chemin optimal pour une rafale donnée. Une stratégie pure prend en considération un seul type d'informations pour déterminer le niveau de congestion dans chaque chemin (tel que le taux de perte ou le délai de bout en bout), et ainsi, sélectionner le chemin optimal. Une stratégie hybride combine plus qu'une stratégie pure pour faire cette sélection. Yang *et al.* [38] concluent qu'en général, les performances du routage multi-chemins sont meilleures que celles du routage à chemin unique (notamment, le routage du plus court chemin).

Ishii *et al.* [39] proposent d'assigner des priorités aux chemins menant d'un nœud source donné à un nœud destination donné. Ainsi, chaque rafale est envoyée sur le chemin qui a la priorité la plus élevée. Par la suite, à chaque fois qu'une rafale est envoyée sur un chemin donné, le nœud source reçoit un message de retour (*feed-back*) qui indique soit le succès, soit l'échec de la transmission de cette rafale. Aussi, les auteurs proposent d'envoyer des paquets de recherche sur les autres chemins menant à la même destination (sur les longueurs d'onde de contrôle) pour recevoir des messages de retour qui indiquent si la transmission de la rafale sur ces différents chemins aurait été réussie ou non. En utilisant les messages de retour, les priorités des chemins sont mises à jour. Nous notons que ces paquets de recherche peuvent produire un surplus (*overhead*) considérable qui peut causer la congestion du plan de contrôle.

Dans [37], Thodime *et al.* proposent un mécanisme d'évitement de la congestion qui sert à sélectionner l'un des chemins (lien-disjoints) entre une source et une destination pour envoyer chaque rafale. Pour cela, chaque nœud du cœur du réseau mesure la charge actuelle du trafic sur chacun de ses liens de sortie et diffuse, *périodiquement*, cette information à tous les nœuds du bord du réseau qui l'utilisent pour éviter les chemins les plus congestionnés lors des envoies de leurs rafales. Aussi, les auteurs proposent un autre mécanisme qui considère une métrique pondérée par le niveau de congestion et la distance (physique ou en termes de sauts) pour améliorer les performances du réseau OBS en termes de la probabilité de perte et du délai de bout en bout des rafales. Notons que le fait de diffuser les informations sur le niveau de congestion périodiquement à tous les nœuds du bord peut occasionner, là aussi, la congestion du plan de contrôle.

Ganguly *et al.* [40] proposent un mécanisme de sondage passif sur les chemins non-optimaux. Dans ce mécanisme, le nœud source envoie une faible fraction de son trafic sur les chemins non-optimaux en gardant très faible la probabilité de les sélectionner. Ensuite, ce nœud source analyse les acquittements positifs (*acknowledgment* [ACK]) et les acquittements négatifs (*negative acknowledgment* [NACK]) durant la période d'une fenêtre de temps coulissante pour déterminer le chemin optimal.

Dans [36], Kiran *et al.* proposent un algorithme pour la sélection des chemins dans un contexte de routage multi-chemins dans les réseaux OBS. Cet algorithme utilise des agents d'apprentissage par renforcement (plus spécifiquement, *Q-learning*) situés dans les nœuds du bord du réseau OBS. Ces agents sélectionnent, pour chaque rafale à transmettre, un chemin optimal en termes de la probabilité de perte des rafales. Après l'envoi de la rafale, le nœud source reçoit un message de retour (*feed-back*) qui est un ACK si la rafale a atteint sa destination avec succès (c'est la destination qui envoie l'ACK dans ce cas) et un NACK si la rafale a été éliminée dans un nœud intermédiaire à cause d'une contention (c'est ce nœud intermédiaire qui envoie le NACK dans ce cas). En utilisant ce message de retour, le nœud source met à jour son appréciation du chemin sur lequel la rafale a été envoyée. Kiran *et al.* [36] reportent que leur approche est plus efficace pour réduire la probabilité de perte comparativement aux travaux de Yang *et al.* [38] et de Ishii *et al.* [39]. Ce travail sera d'intérêt pour nous dans la suite de cette thèse puisque notre travail sur le routage multi-chemins (dans le chapitre 3) sera comparé à celui-ci.

D'autres travaux tels que Lu *et al.* [41], Teng *et al.* [42] et Klinkowski *et al.* [43] s'intéressent, essentiellement, à la détermination de la liste des chemins optimaux pour chaque paire de nœuds (source, destination) en utilisant des techniques d'optimisation combinatoire. De plus, Teng *et al.* [42] poursuivent une approche basée sur l'ingénierie du trafic pour atteindre cet objectif. Aussi, dans les solutions proposées par Teng *et al.* [42] et par Klinkowski *et al.* [43] une entité centrale prend les décisions pour le routage multi-chemins.

Li *et al.*, dans [44], considèrent le routage multi-chemins dans le cas spécifique de deux chemins prédéterminés entre chaque paire de nœuds (source, destination). L'idée est de distribuer la charge du trafic entre la paire de nœuds sur les deux chemins en prenant en considération le niveau de congestion dans chaque chemin ainsi que sa taille en termes du nombre de sauts (et ainsi, en termes du délai de bout en bout). Pour obtenir le niveau de congestion de chaque chemin, des paquets de sondage sont utilisés.

Hirota *et al.* [45] et Argos *et al.* [46] se distinguent par leur proposition de mécanismes distribués où chaque nœud (incluant les nœuds intermédiaires) décide sur quel lien de sortie une rafale doit être envoyée pour atteindre sa destination. Par exemple, dans [45], chaque nœud possède une table où pour chaque lien de sortie et chaque longueur d'onde est stockée une valeur, appelée *Suitability Index (SI)*, qui représente l'estimation de la qualité d'une longueur d'onde sur un lien de sortie donné. Les SIs sont mises à jour via des messages de retour dans les deux cas d'une transmission réussie ou échouée d'une rafale.

Le Tableau 2 présente un récapitulatif des travaux de recherche sur le routage multi-chemins dans les réseaux OBS. Nous constatons que la majorité des travaux déploient leurs solutions dans les nœuds source, utilisent un calcul statique des chemins alternatifs et considèrent la probabilité de perte comme métrique unique pour évaluer la qualité des chemins. Aussi, la plupart des travaux utilisent des messages de retour explicites de bout en bout pour évaluer la qualité des chemins, ce qui peut congestionner sévèrement le plan de contrôle. Ainsi, il sera intéressant d'explorer une solution pour le routage multi-chemins qui soit déployée sur tous les nœuds du réseau OBS (déploiement distribué), qui minimise le surplus de communication, qui considère en même temps la probabilité de perte et le délai de bout en bout comme métriques de sélection des chemins et pour laquelle le calcul des chemins du routage est dynamique selon l'état actuel du réseau. Cette solution sera présentée dans le chapitre 3.

Tableau 2. Récapitulatif des travaux sur le routage multi-chemins dans les réseaux OBS.

Travaux	**Déploiement**	**Calcul des chemins**	**Sélection des chemins**	**Métriques**
Yang *et al.* [38]	Nœuds source	Statique	Classement des routes	Perte
Ishii *et al.* [39]	Nœuds source	Statique	Classement des routes	Perte
Ganguly *et al.* [40]	Nœuds source	Statique	Classement des routes	Perte
Kiran *et al.* [36]	Nœuds source	Statique	Classement des routes	Perte
Thodime *et al* [37]	Nœuds source	Statique	Classement des routes	Perte et délai
Li *et al.* [44]	Nœuds source	Statique	Probabiliste	Perte et délai
Lu *et al*. [41]	Nœuds source	Statique (optimisé)	Probabiliste	Perte
Teng *et al.* [42]	Centralisé	Statique (optimisé)	Probabiliste	Perte
Klinkowski *et al.* [43]	Centralisé	Statique (optimisé)	Probabiliste	Perte
Argos *et al.* [46]	Distribué	Statique	Probabiliste	Perte
Hirota *et al.* [45]	Distribué	Statique	Classement des liens	Perte

2.3. Le routage alternatif

Dans le routage alternatif nous nous intéressons au routage par déflection comme approche de résolution des contentions. En effet, bien que les approches proactives, tel que le routage multi-chemins, améliorent les performances des réseaux OBS, les approches réactives (décrites dans la section 2.1.5) sont indispensables pour résoudre les contentions après leurs apparitions.

Le routage par déflection a été identifié, dès l’apparition des réseaux OBS, comme étant une approche importante pour la résolution des contentions. Cependant, les premiers travaux dans ce domaine se sont focalisés sur des réseaux avec des topologies régulières et des modèles de trafic particuliers [47].

Wang *et al.* [48], proposent un protocole de routage par déflection qui utilise deux fonctionnalités : *sender check* et *sender retransmission*. Ces deux fonctionnalités servent, respectivement, à empêcher la déflection d'une rafale à son nœud source et à décider si la rafale doit être retransmise depuis son nœud source. Ces fonctionnalités visent à contrôler le routage par déflection.

Dans [49], Hsu *et al.* étudient les performances des réseaux OBS qui utilisent le protocole de réservation JET (cf. section 2.1.4) quand le routage par déflection est adopté conjointement avec les FDLs comme des approches de résolution des contentions. La conclusion générale énoncée dans cette étude est que le routage par déflection apporte une amélioration des performances quand le nombre de longueurs d'onde est relativement petit et quand la charge du trafic est modérée.

Chen *et al.*, dans [50], étudient la probabilité de perte des rafales dans un nœud du cœur du réseau OBS qui utilise le routage par déflection pour résoudre les contentions. Les deux cas avec et sans convertisseurs de longueur d'onde sont considérés. Selon les auteurs, contrairement à d'autres modèles qui ne considèrent qu'un seul port de déflection, leur modèle peut être appliqué à un nœud OBS du cœur avec n'importe quel nombre de ports de déflection.

Dans [51], Zalesky *et al.* proposent une approche de réservation des longueurs d'onde pour le routage par déflection. Cette approche s'inspire de l'approche de réservation des ressources, dite *trunk reservation*, utilisée dans les réseaux optiques à commutation de circuits. Cette approche contrôle la déflection en imposant une limite sur la quantité de trafic routé par déflection quand la charge du trafic dans le réseau est élevée. Cet objectif est atteint en réservant un nombre de longueurs d'onde à l'usage exclusif des rafales non routées par déflection.

Dans [52], Cameron *et al.* proposent une approche appelée SP-PRDR (*Shortest Path - Prioritized Random Deflection Routing*). Dans SP-PRDR, la déflection est effectuée vers un lien de sortie aléatoire en cas de contention. La rafale routée par déflection se voit assigner une priorité inférieure à la priorité par défaut des rafales non routées par déflection. Ainsi, en cas de contention entre une rafale routée par déflection et une autre rafale non routée par déflection, cette dernière est privilégiée et la première est éliminée ou sa réservation est annulée.

Dans [53], Lee *et al.* proposent un mécanisme appelé *Contention-based Limited Deflection Routing* (CLDR). Dans CLDR, les deux fonctionnalités suivantes sont effectuées : (a) déterminer si une rafale doit être routée par déflection ou retransmise depuis son nœud source après l'apparition d'une contention; (b) si la décision est de router la rafale par déflection, alors le lien de déflection optimal, qui minimise le délai et la probabilité de perte de la rafale est sélectionné. La deuxième fonctionnalité (b) est basée sur un modèle d'optimisation qui doit être résolu périodiquement.

Lee *et al.*, Dans [54], proposent un modèle d'optimisation qui détermine, dynamiquement, le chemin de déflection qui a la moindre charge de trafic. De plus, les rafales routées par déflection qui possèdent une priorité élevée, se voient assigner des longueurs d'onde virtuellement sur le chemin de déflection, c.-à-d., qu'elles peuvent utiliser n'importe quelle longueur d'onde disponible sur le chemin de déflection en utilisant la conversion de longueur d'onde.

Coutelen *et al.*, dans [55], s'intéressent au problème de la sélection du lien de déflection. Pour cela, trois stratégies sont proposées, à savoir, *Least Loaded Node* (LLN), *LLN/U-turn* et *LLN/U-turn/ Load Balancing*. LNN consiste à envoyer la rafale à router par déflection sur le lien de sortie qui mène vers le voisin le moins congestionné en excluant le nœud qui a déjà envoyé la rafale. L'option *U-turn* (demi-tour) combinée avec LLN consiste à envoyer la rafale vers le nœud qui l'a déjà envoyé si elle ne peut être envoyée à aucun autre nœud voisin. À noter ici que les auteurs supposent que tous les liens sont bidirectionnelles et que chaque lien est composé de deux fibres chacune utilisée dans un seul sens. Avec l'option *Load Balancing* (LB) un équilibrage de charge est appliqué sur le trafic dans le réseau [56]. Les auteurs reportent que les options U-turn et LB combinées à la stratégie LLN apportent une réduction considérable de la probabilité de perte des rafales sans pour autant augmenter leur délai moyen de bout en bout de façon significative.

Dans [57], Ogino *et al.* proposent un mécanisme de routage par déflection qui considère la possibilité de contention dans les nœuds en aval d'un nœud du cœur. Une métrique appelée la distance prévue de la route vers la destination (*expected route distance to destination*) est définie. Cette métrique considère la possibilité de contention à chaque nœud de transit en aval. Elle est calculée en utilisant les taux de perte mesurés dans les liens de sortie des nœuds en aval et elle est utilisée pour sélectionner le lien de déflection en cas de contention.

Du *et al.*, Dans [58], proposent un algorithme de routage par déflection basé sur l'équilibrage de la charge du trafic dans le réseau OBS. Cet algorithme, appelé *Load Balancing Deflection Routing algorithm* (LBDR), prend en considération la route primaire, la topologie du réseau, ainsi que la distribution de la charge du trafic dans le réseau pour calculer le chemin de déflection. LBDR est centralisé et nécessite une quantité considérable d'informations pour calculer les chemins de déflection.

Vokkarane *et al.* [59] proposent des politiques basées sur la combinaison du routage par déflection et de la segmentation. Ainsi, selon la priorité des rafales en contention et la taille de la partie impliquée dans la contention de chaque rafale, la décision est prise sur le routage par déflection, la segmentation ou encore, l'élimination de la rafale selon la politique adoptée.

Le Tableau 3 montre un récapitulatif des travaux sur le routage alternatif dans les réseaux OBS. Nous constatons que le calcul des chemins alternatifs est statique et que la sélection des chemins alternatifs est fixe dans la plupart des travaux sur le routage alternatif dans les réseaux OBS. Dans le chapitre 3, nous présenterons une solution pour le routage alternatif où le calcul des chemins est dynamique et la sélection des chemins est basée sur l'état actuel du réseau OBS.

Tableau 3. Récapitulatif des travaux sur le routage alternatif dans les réseaux OBS.

Travaux	**Adaptation**	**Déploiement**	**Calcul des chemins**	**Sélection des chemins**
Kim *et al.* [47]	Non	Distribué	Statique	Fixe
Wang *et al.* [48]	Non	Distribué	Statique	Fixe
Hsu *et al.* [49]	Non	Distribué	Statique	Fixe
Chen *et al.* [50]	Non	Distribué	Statique	Probabiliste
Zalesky *et al* [51]	Non	Distribué	Statique	Fixe
Cameron *et al.* [52]	Non	Distribué	Statique	Fixe
Vokkarane *et al.* [59]	Non	Distribué	Statique	Fixe
Du *et al.* [58]	Non	Centralisé	Statique	Fixe
Lee *et al* [53]	Oui	Centralisé/distribué	Statique(optimisé)	Avec seuils
Lee *et al.* [54]	Oui	Centralisé/distribué	Statique(optimisé)	Avec seuils
Coutelen *et al.*[55]	Oui	Distribué	Statique	Classement
Ogino *et al.* [57]	Oui	Distribué	Statique	Classement

2.4. La sélection des longueurs d'onde

Bien que la sélection des longueurs d'onde possède un rôle clé dans la transmission des rafales dans les réseaux OBS, elle n'a pas reçu beaucoup d'attention dans les travaux de recherche sur les réseaux OBS. Aussi, la majorité des travaux de recherche qui s'intéressent aux réseaux OBS supposent que la conversion de longueur d'onde est omniprésente. Néanmoins, comme déjà souligné dans la section 2.1.5, cette hypothèse est simpliste à cause des enjeux économiques et technologiques qui entourent l'utilisation des convertisseurs de longueur d'onde dans les réseaux OBS. Ainsi, cette hypothèse pourrait surestimer les performances du réseau OBS dans un scénario de déploiement réel. En effet, le déploiement des convertisseurs de longueur d'onde dans chaque nœud du réseau OBS éliminera complètement *la contrainte de continuité de longueur d'onde* (*wavelength continuity constraint*) qui stipule que chaque rafale doit être transmise sur la même longueur d'onde de sa source à sa destination. Dans ce cas, une rafale peut changer de longueur d'onde dans n'importe quel nœud intermédiaire pour éviter son élimination à cause d'une contention. En outre, dans la plupart des travaux sur les réseaux OBS, les auteurs adoptent des politiques classiques de réservation des ressources, tels que les politiques *First-Fit*, *Least Used* ou *Random*. Ces politiques sont habituellement utilisées dans les réseaux optiques à commutation de circuits pour assigner une longueur d'onde à chaque connexion [60]. Cependant, Teng *et al.* [61] ont montré que ces politiques ne sont pas adaptées à la sélection des longueurs d'onde dans les réseaux OBS. À noter que, la plupart des algorithmes de sélection des longueurs d'onde proposés pour les réseaux OBS sont adaptatifs et basés sur les messages de retour [36, 61, 62].

Parmi les algorithmes de sélection des longueurs d'onde pour les réseaux OBS, nous retrouvons l'algorithme appelé *Latest Available Unscheduled Channel with Void Filling* (LAUC-VF) qui est proposé par Xiong *et al.* [32]. LAUC-VF combine une politique classique d'allocation des ressources dite *Latest Available Unscheduled Channel* (LAUC) avec la technique dite *Void Filling* (VF). LAUC consiste à utiliser la longueur d'onde qui est devenue disponible le plus récemment dans le temps. VF consiste à exploiter les vides entre les réservations successives des rafales pour ordonnancer une nouvelle rafale. Le but de LAUC-VF est de réduire la fragmentation et d'améliorer l'utilisation des longueurs d'onde.

Dans [62], Wang *et al.* proposent un algorithme, appelé *Priority-based Wavelength Assignment* (PWA). Dans PWA, chaque nœud dans le réseau OBS maintient localement une priorité pour chaque paire (longueur d'onde, nœud de destination). Les priorités sont mises à jour en utilisant les messages de retour reçus après la transmission d'une rafale de données. Ainsi, si la rafale a atteint sa destination, la priorité de la paire (longueur d'onde sur laquelle la rafale a été envoyée, destination de la rafale) est augmentée. Sinon, cette priorité est diminuée. Pour envoyer une rafale de données, le nœud source sélectionne la longueur d'onde libre dont la priorité est la plus élevée. Wang *et al.* rapportent dans [62] que les performances de PWA sont meilleures que celle de la politique de sélection aléatoire (*Random*) en termes de la probabilité de perte des rafales. Cependant, cet avantage diminue pour devenir marginal lorsque la charge du trafic devient très élevée.

Teng *et al.* [61] proposent deux variantes de PWA qui sont *PWA-link* et *PWA-lambda*. Dans PWA-link, chaque nœud associe une priorité à chaque paire (longueur d'onde, lien) au lieu de la paire (longueur d'onde, nœud de destination) dans PWA. Ainsi PWA-link possède une granularité plus fine que celle de PWA au détriment de la quantité d'informations échangées dans le réseau OBS. Dans PWA-lambda, chaque nœud associe une priorité à chaque longueur d'onde. Ceci rend PWA-lambda simple à implanter mais avec une dégradation des performances comparativement à PWA et PWA-link. En effet, les résultats numériques dans [61] montrent, qu' en termes de la probabilité de perte des rafales, PWA-link est le meilleur et PWA-lambda est le pire lorsque PWA, PWA-link et PWA-lambda sont comparés. Aussi, Teng *et al.* [61] proposent une approche non-adaptative de sélection des longueurs d'onde, appelée *First-Fit-TE* où TE est l'acronyme de *Traffic Engineering*. Dans First-Fit-TE, une longueur d'onde, dite *de début*, est assignée à chaque nœud. Cette assignation prend en considération le modèle du trafic dans le réseau et la configuration *fixe* des chemins du routage entre les nœuds du réseau. Quand First-Fit-TE est utilisé, chaque nœud cherche une longueur d'onde disponible, pour envoyer une rafale, en commençant de sa longueur d'onde de début. La raison derrière cette façon de chercher une longueur d'onde disponible est d'améliorer la façon de la politique *First-Fit* à chercher une longueur d'onde disponible en essayant d'isoler les trafics des différents nœuds dans le réseau OBS. Cependant, le fonctionnement de First-Fit-TE dépend fortement du modèle de trafic *prédéterminé* et de la configuration *fixe* des chemins du routage, ce qui peut limiter son utilisation quand le modèle du trafic est dynamique (ou même incertain) ou quand le routage multi-chemins est adopté.

Kiran *et al.* [36] proposent un algorithme de sélection des longueurs d'onde, appelé *Q-learning algorithm for Wavelength Selection* (QWS). QWS est basé sur les messages de retour et il est différent de PWA dans la manière de mettre à jour les priorités des paires (longueur d'onde, nœud de destination). En effet, QWS adopte une approche d'apprentissage par renforcement pour effectuer ces mises à jour. Selon Kiran *et al.* [36], QWS est meilleur que PWA-link en termes de la probabilité de perte dans le réseau OBS. QWS sera d'intérêt pour nous dans cette thèse puisque nous comparerons ses performances directement aux performances de l'algorithme de sélection des longueurs d'onde BETWA qui sera présenté dans le chapitre 4 et indirectement aux performances de l'algorithme de sélection des longueurs d'onde PBEWS, qui sera présenté dans le chapitre 5.

2.5. Le provisionnement absolu de la qualité de service

Dans la littérature sur les réseaux OBS, il existe deux modèles de provisionnement de la Qualité de Service (QdS) : le modèle *relatif* [63-73] et le modèle *absolu* [74-81].

Dans le modèle relatif, les performances d'une classe de trafic d'une priorité donnée sont garanties d'être meilleures qu'une autre classe de trafic de priorité moindre sans pour autant garantir quantitativement le niveau de performance de chacune des deux classes. Les techniques principales utilisées dans le modèle relatif sont : le temps de décalage [63], la segmentation [64, 65], la préemption probabiliste [66, 67], l'ordonnancement des paquets de contrôle [68, 69], la différentiation proportionnelle de la QdS [70], l'allocation des tampons [71, 72] et l'ordonnancement des rafales [72, 73]. Des aperçus sur les travaux qui s'intéressent au modèle relatif de provisionnement de la QdS dans les réseaux OBS peuvent être trouvés dans [82] et [83].

Dans le modèle absolu, contrairement au modèle relatif, les performances d'une classe de trafic de priorité élevée sont garanties quantitativement en utilisant des seuils prédéfinis qui définissent une performance *pire-cas* pour chaque métrique de performance considérée. Par exemple, un mécanisme de provisionnement de la QdS qui garantit que le taux de perte d'une classe de trafic donnée ne dépasse jamais 0.1 peut être classé comme étant un mécanisme de QdS absolu. Bien que le modèle absolu est, généralement, plus complexe à implanter que le modèle relatif [82], il est préférable du point de vue des applications et des utilisateurs. En effet, avec ce modèle, les utilisateurs obtiennent des garanties fermes sur les performances de leur trafic. Ceci est particulièrement utile dans le cas des applications sensibles aux pertes et aux délais. Il est à noter que la métrique principale de QdS à l'intérieur du réseau OBS est la probabilité de perte des rafales puisque les rafales sont envoyées dans le réseau OBS sans subir des délais de stockage additionnels, notamment en l'absence des lignes de retardement (FDLs).

Les travaux ayant proposés des solutions pour le provisionnement absolu de la QdS sont basés, essentiellement, sur : l'élimination anticipée (*Early Dropping* [ED]) [74], le groupage des longueurs d'onde (*Wavelength Grouping* [WG]) [74, 75, 77, 79] et la préemption [77, 80]. À noter ici que dans certains travaux, ces techniques sont combinées.

Zhang *et al.*, dans [74], proposent deux mécanismes pour le provisionnement absolu de QdS : *Early Dropping* (ED) et *Wavelength Grouping* (WG). *Early Dropping* élimine de façon anticipée et probabiliste les rafales de priorité inférieure afin de maintenir le niveau requis de probabilité de perte des rafales de priorité supérieure. À noter que les classes de trafic sont assignées des priorités selon leurs exigences en termes de probabilité de perte. Ainsi, chaque nœud surveille les taux de perte de chaque classe de trafic. Pour décider si une rafale appartenant à une classe de trafic doit être éliminée ou non, une probabilité, appelée *early dropping probability*, est utilisée. Pour une classe *i*, cette probabilité, notée

$P_{C_i}^{ED}$, est calculée selon le taux de perte observé et le seuil de perte toléré de la classe supérieure suivante $i+1$. Selon $P_{C_i}^{ED}$, un drapeau d'élimination (*dropping flag*), noté e_i, détermine si la rafale arrivée qui appartient à la classe *i* doit être éliminée ou non. Ceci est effectué en tirant un nombre aléatoire entre 0 et 1. Si ce nombre est inférieur à $P_{C_i}^{ED}$, e_i prend la valeur 1, c.-à-d., la rafale doit être éliminée. Sinon, e_i prend la valeur 0. De plus, si la valeur d'au moins un des drapeaux des classes supérieures (e_{i+1}, e_{i+2}, … e_{N-1}) est 1, la rafale de la classe *i* doit être éliminée. L'inconvénient de ED provient du fait que certaines rafales peuvent être éliminées sans qu'ils auraient, nécessairement, causé des contentions et des pertes, ce qui se traduit en un gaspillage de la bande passante. Cet inconvénient provient de l'aspect probabiliste de ED. *Wavelength Grouping* alloue un minimum de longueurs d'onde à chaque classe de trafic selon ses exigences en termes de probabilité de perte. Pour cela, la fameuse formule *Erlang-B* [84] est utilisée pour déterminer le nombre de longueurs d'onde nécessaire pour une classe de trafic selon sa charge dans le réseau et selon son seuil maximal de probabilité de perte tolérée. Il existe deux variantes de WG : *Static Wavlength Grouping* (SWG) et *Dynamic Wavelength Grouping* (DWG). Dans SWG, les longueurs d'onde sont assignées de manière fixe à chaque classe de trafic, c.-à-d., les rafales d'une classe de trafic ne peuvent utiliser que les longueurs d'onde qui ont été assignées à leur classe. Dans DWG, les rafales peuvent utiliser n'importe quelle longueur d'onde, pourvue que le nombre total de longueurs d'onde utilisées par les rafales de cette classe ne dépasse pas le nombre de longueurs d'onde qui lui a été déterminé. L'inconvénient des mécanismes basés sur WG est la diminution du multiplexage statistique et l'utilisation inefficace des longueurs d'onde dans certains scénarios. En effet, dans le cas de SWG, une rafale appartenant à une classe de trafic donnée peut être éliminée même s'il existe des longueurs d'onde disponibles parmi les longueurs d'onde assignées aux autres classes de trafic.

Dans un contexte similaire à *Wavelength Grouping*, Kim *et al.* [75] proposent un mécanisme, appelé *Dynamic Virtual Lambda Partitioning* (DVLP). DVLP fonctionne au niveau de chaque lien dans le réseau OBS. Il consiste à allouer, initialement, dans la phase de conception du réseau, un nombre de longueurs d'onde à chaque classe de trafic. Cette allocation initiale est basée sur le niveau de QdS requis de chaque classe et une estimation de sa charge de trafic dans le réseau. Les auteurs supposent qu'au moins une longueur d'onde est allouée à chaque classe de trafic. Notons ici qu'une classe de trafic de priorité *i* peut utiliser non seulement ses longueurs d'onde mais aussi toutes les longueurs d'onde des classes de trafic de priorité inférieure (*i-1*, *i-2*, … 1). Par la suite, à chaque fois que le taux de perte d'une classe de trafic dépasse son seuil maximal de probabilité de perte, une longueur d'onde est retirée de la classe 1 (*best effort*) et allouée à cette classe. Cependant, pour éviter le problème de famine (*starvation*) pour le trafic *best effort*, une classe de trafic ne peut avoir plus de longueurs d'onde qu'un nombre maximal prédéfini. Par ailleurs, pour éviter le problème de fluctuation dans le processus de reconfiguration de l'allocation des longueurs d'onde, un mécanisme basé sur des seuils est développé. Kim *et al.* [75] affirment que DVLP ne peut garantir le provisionnement absolu de la qualité de service dans tous les cas. En effet, ceci peut dépendre de plusieurs facteurs, tels quel le nombre de longueurs d'onde dans chaque fibre, le nombre de classes de trafic et la proportion et la charge de chaque classe de trafic. Pour cela, une analyse a été faite pour la faisabilité de DVLP dans un scénario simple qui ne comporte que deux classes de trafic.

Yang *et al.* [81] propose une série de politiques pour le partage des ressources entre plusieurs classes de trafic. Ces politiques sont des variantes du mécanisme *Dynamic Wavelength Grouping* proposé dans [74]. En effet, en variant le nombre minimum et le nombre maximum de longueurs d'onde qu'une classe de trafic peut occuper, les auteurs définissent les politiques suivantes : *Wavelength Partitioning* (WP) qui représente DWG tel que présenté dans [74], *Wavelength Sharing with Maximum occupancy* (WS-Max), *Wavelength Sharing with Minimum provisioning* (WS-Min) et *Wavelength Sharing with Minimum provisioning and Maximum occupancy* (WS-MinMax). Les auteurs reportent que WS-MinMax réduit la probabilité de perte du trafic *best effort* ainsi que la probabilité de perte totale (de toutes les classes) comparativement à WP.

Le mécanisme proposé par Phuritatkul *et al.* [76] garantit une probabilité de perte maximale pour chaque classe de trafic exigeant une garantie. Pour cela, chaque nœud du cœur collecte des statistiques sur chaque classe de trafic. Une rafale garantie peut préempter une rafale non-garantie (c.-à-d., annuler sa réservation) proactivement, si le taux de perte de la classe de trafic de la rafale garantie s'approche de son seuil maximal de probabilité de perte. Dans ce cadre, une probabilité de préemption $W(n, h)$ est définie pour chaque classe garantie n et chaque lien h. $W(n, h)$ prend la valeur 0, c.-à-d., la préemption est désactivée si le taux de perte des rafales de la classe n sur le lien h est inférieur à un seuil P_{min}. Il prend la valeur 1, c.-à-d., la préemption est activée, si le taux de perte des rafales de la classe n sur le lien h est supérieur au seuil maximal de probabilité de perte $P_{n,h}$. Cependant, lorsque le taux de perte est entre P_{min} et $P_{n,h}$, $W(n, h)$ prend une valeur progressive entre 0 et 1. En outre, les auteurs ont développé un modèle markovien à temps continu pour analyser la probabilité de perte dans un lien OBS quand le mécanisme de préemption proposé est adopté. La préemption est une solution plus élégante et plus efficace que l'élimination anticipée puisqu'elle n'élimine les rafales que lorsque c'est nécessaire. Néanmoins, la préemption peut causer un gaspillage de la bande passante lorsqu'une rafale est préemptée dans un nœud alors que son paquet de contrôle avait déjà réservé des ressources dans les nœuds en aval vers sa destination. Il s'agit dans ce qui est appelé une rafale *fantôme*.

Guan *et al.* [77] proposent de combiner la préemption avec un mécanisme de réservation virtuelle des longueurs d'onde, appelé *Virtual Channel Reservation* (VCR). Dans VCR, k_i longueurs d'onde (où $0 \leq k_i \leq T$ et T est le nombre de longueurs d'onde dans chaque fibre) sont allouées à chaque classe de trafic i. Le nombre k_i est défini en utilisant la formule *Erlang-B* et en prenant en considération le seuil maximal de probabilité de perte de chaque classe de trafic. Si des longueurs d'onde sont disponibles, une rafale appartenant à la classe i peut réserver n'importe quelle longueur d'onde même si l'utilisation des longueurs d'onde de la classe i dépasse k_i. Néanmoins, si cette rafale trouve toutes les longueurs d'onde occupées, elle peut préempter une rafale de priorité inférieure seulement si l'utilisation des longueurs d'onde de sa classe i ne dépasse pas k_i. Ce fonctionnement permet à toutes les classes de trafic d'utiliser pleinement la bande passante de la fibre optique tout en garantissant le niveau de QdS requis par chaque classe. Aussi, pour remédier au problème des rafales fantômes, et ainsi, améliorer l'utilisation de la bande passante, Guan *et al.* [77] proposent d'envoyer un paquet de contrôle qui a pour rôle d'annuler les réservations des ressources de la rafale preemptée dans les nœuds en aval.

Phùng *et al.* [78] proposent un *framework* pour garantir le provisionnement absolu de la QdS à chaque lien du réseau OBS et au niveau de chaque flot de bout en bout. À cette fin, la préemption est adoptée conjointent avec un mécanisme de contrôle d'admission. Le rôle de la préemption est de garantir que le seuil maximal de probabilité de perte de chaque classe de trafic est respecté. Cet objectif est réalisé en portant les pertes d'une classe de trafic dont le taux de perte approche de son seuil maximal à d'autres classes dont les taux de perte sont loin de leurs seuils maximaux. Dans ce sens, même une rafale de priorité inférieure peut préempter une rafale de priorité supérieure si le taux de perte de la première approche de son seuil maximal alors que le taux de probabilité de la deuxième est loin de son seuil maximal. Par ailleurs, le rôle du mécanisme de contrôle d'admission est la réduction de la charge du trafic dans chaque lien du réseau afin de permettre que les exigences de toutes les classes de trafic, en termes de la probabilité de perte, soient satisfaites. En outre, Phùng *et al.* [78] adoptent une politique de préemption qui permet d'atteindre l'équité (*fairness*) *intra-classe* au sein d'une classe de trafic. En effet, dans les autres travaux sur le provisionnement absolu de la QdS, même si le seuil maximal de perte d'une classe donnée *i* est respecté, certains flots appartenant à la classe *i* peuvent souffrir d'un taux de perte qui dépasse le seuil de la classe. Phùng *et al.* [78] résolvent ce problème en rendant *probabiliste* le choix de préempter une rafale parmi un ensemble de rafales appartenant à la même classe de trafic. Aussi, les auteurs tiennent en compte la durée de réservation des rafales pour prendre en considération le fait que les rafales qui ont des réservations de longues durées subissent, généralement, plus de préemptions que les rafales qui ont des réservations de courtes durées.

Hongbo *et al.* [79] proposent le mécanisme *Reserve-and-Preempt Scheme* (RPS) pour améliorer le provisionnement de la bande passante du trafic *best-effort* dans le contexte du provisionnement absolu de la QdS. Dans RPS, une métrique, appelée *Distance To Threshold* (DTT), mesure la différence entre le taux de perte actuel d'une classe de trafic garantie et son seuil maximal de probabilité de perte. Selon la valeur du DTT et en utilisant un scénario parmi plusieurs proposés, il peut arriver qu'une rafale *best-effort* préempte une rafale garantie si le DTT de la classe de trafic de la rafale garantie est : (a) assez grand; et (b) le maximum parmi les autres DTTs des classes de trafic garanties. Les même auteurs, dans [80], proposent plusieurs mécanismes qui combinent RPS et le groupage des longueurs d'onde *(wavelength Grouping)* pour assurer que les classes de trafic garanties peuvent être servies avec leurs seuils maximaux de probabilité de perte.

Le Tableau 4 présente un récapitulatif des travaux sur le provisionnement absolu de la QdS du point de vue de la technique utilisée (élimination anticipée, groupage des longueurs d'onde et préemption), la granularité de la solution (au niveau d'un lien, d'un nœud ou encore d'un chemin de bout en bout), l'amélioration des performances du trafic *best effort*, la solution au problème des rafales fantômes dans le cas où la préemption est adoptée ainsi que l'équité intra-classe de la solution. Nous constatons que le travail de Phùng *et al.* [78] se distingue par sa considération à la fois du provisionnement absolu de la QdS au niveau de chaque lien ainsi qu'au niveau d'un flot de bout en bout. Aussi, ce même travail se distingue par sa solution au problème de l'équité intra-classe. Par ailleurs, nous constatons que les travaux antérieurs sur le provisionnement absolu de la QdS considèrent un seuil maximal de probabilité de perte pour chaque classe de trafic, et ce, même pour les classes de trafic qui sont les plus sensibles aux pertes et qui ont la priorité la plus élevée. En effet, dans toutes les propositions antérieures, les contentions intra-classes (entre des rafales appartenant à la même classe de trafic) peuvent surgir dans le cœur du réseau OBS, ce qui se traduit par un taux de perte qui peut être non négligeable pour les rafales des classes de priorités élevées. Dans cette thèse, nous nous intéressons à une problématique plus avancée dans le contexte du provisionnement absolu de la QdS qui consiste à garantir une transmission *sans pertes* aux rafales appartenant aux classes sensibles aux pertes à l'intérieur du réseau OBS. Aussi, nous nous intéressons à l'amélioration des performances du trafic *best effort* en termes de la probabilité de perte.

Tableau 4. Récapitulatif des travaux sur le provisionnement absolu de la QdS dans les réseaux OBS.

Travaux	**Élimination anticipée**	**Groupage des longueurs d'onde**	**Préemption**	**Granularité**		**Performance du trafic *best effort***	**Rafales fantômes**	**Équité *intra-classe***
				Lien / nœud	**De bout en bout**			
Zhang *et al.* [74]	√	√		√				
Kim *et al.* [75]		√		√				
Yang *et al.* [81]		√		√	√			
Phuritatkul *et al.* [76]			√	√				
Guan *et al.* [77]		√	√	√			√	
Phùng *et al.* [78]			√	√	√		√	√
Hongbo *et al.* [79]			√	√		√		
Hongbo *et al.* [80]		√	√	√		√		

Chapitre 3 :

Novel reinforcement learning-based approaches to reduce loss probability in buffer-less OBS networks

Abdeltouab Belbekkouche, Abdelhakim Hafid, Michel Gendreau

Abstract

Optical Burst Switching (OBS) is a promising switching paradigm for the next generation Internet. A buffer-less OBS network can be implemented simply and cost-effectively without the need for either wavelength converters or optical buffers which are, currently, neither cost-effective nor technologically mature. However, this type of OBS networks suffers from relatively high loss probability caused by wavelength contentions at core nodes. This could prevent or, at least, delay the adoption of OBS networks as a solution for the next generation optical Internet. To enhance the performance of buffer-less OBS networks, we propose three approaches: (a) a reactive approach, called Reinforcement Learning-Based Deflection Routing Scheme (RLDRS) that aims to resolve wavelength contentions, after they occur, using deflection routing; (b) a proactive multi-path approach, called Reinforcement Learning-Based Multi-path Routing (RLMR), that aims to reduce wavelength contentions; and (c) an approach, called Integrated Reinforcement Learning-based Routing and Contention Resolution (IRLRCR), that combines RLMR and RLDRS to conjointly deal with wavelength contentions proactively and reactively. Simulation results show that both RLMR and RLDRS reduce, effectively, loss probability in buffer-less OBS networks and outperform the existing multi-path and deflection routing approaches, respectively. Moreover, simulation results show that a substantial performance improvement, in terms of loss probability, is obtained using IRLRCR.

Keywords: Multi-path routing; Deflection Routing; Optical Burst Switching; Unsupervised learning; Wavelength Division Multiplexing

3.1. Introduction

WDM (Wavelength Division Multiplexing) is an attractive technology to support the huge amount of data required by the future optical Internet. It uses the potential capacity in optical fibers that contains many wavelengths able to carry many Gbps by using statistical multiplexing. This potential requires good switching technology to efficiently exploit it. OBS (Optical Burst Switching) [18, 19] is a good switching paradigm candidate to fill this need. It has received an increasing interest from researchers over the last several years, since it presents a good tradeoff between traditional Optical Circuit Switching (OCS), which is relatively easy to implement but suffers from poor bandwidth utilization and coarse granularity, and Optical Packet Switching (OPS) which has a good bandwidth utilization and fine granularity but suffers from complex implementations because of the immaturity of the current technologies, such as optical buffers and ultra fast optical switches [18].

In OBS networks, data packets with the same destination are aggregated in bursts of variable lengths at the ingress node, this is called Burst Assembly. After burst assembly, a Control Packet (called also Burst Header Packet) is sent, using a dedicated control wavelength, from source to destination in order to reserve the required resources along a lightpath. This control packet is subject to Optical-Electronic-Optical (OEO) conversions at each core node (OBS switch) where it receives an appropriate processing. After a delay called Offset Time (OT), the corresponding data burst is sent, on one of the data wavelengths, through the same lightpath without any buffering requirement inside the OBS network.

A major issue in OBS networks is wavelength contention which is the main cause of burst losses; this may result in a high burst loss probability (defined as the rate of bursts lost in the OBS network core nodes) causing a considerable performance degradation. A contention arises when two or more bursts intend to take the same output fiber, on the same wavelength, at the same time. There exist mainly two kinds of approaches to deal with wavelength contention: reactive approaches and proactive approaches. Whereas reactive approaches try to resolve contentions after they occur in a network core node (generally, based on local information of this node), proactive approaches attempt to prevent contentions, and consequently burst losses, from occurring.

Proactive approaches are, generally, implemented at OBS network ingress nodes by sending a data burst on the optimal path towards a given destination (in the context of multi-path routing) and/or assigning an optimal wavelength to it [36, 85]. Optimality of a path or a wavelength can be measured by the level of burst loss probability, burst end-to-end delay or a metric that combines both [37, 85].

Reactive approaches include: (a) buffering: send one of the contended bursts on its primary output port and buffer the other bursts using Fiber Delay Lines (FDLs); (b) wavelength conversion: by sending one of the contended bursts on its initial wavelength and the other bursts on an alternative available wavelength; (c) burst segmentation: where the contending part of a burst involved in a contention is dropped, forwarded on an alternate path or buffered while the other part is forwarded on the primary path; and (d) deflection routing: where only one of the bursts involved in a contention is routed to its primary output fiber, whereas each of the other bursts is switched to an alternate outgoing fiber with available wavelengths. We note that combining one or more of these techniques often leads to better improvement in the OBS networks [86].

In this paper, we consider a buffer-less OBS network without FDLs and without wavelength converters which make the network cost-effective. We adopt Just Enough Time (JET) [18] protocol for resource reservation. We propose three novel schemes to reduce the loss probability of the buffer-less OBS network: (1) a novel deflection routing (reactive) scheme, called Reinforcement Learning-based Deflection Routing Scheme (RLDRS) [87], to reduce loss probability when wavelength contentions occur. RLDRS is a distributed scheme; it is concerned with the selection of the alternative route in case of deflection. Each node in the network has to learn the best alternative output link to deflect bursts for a given destination based on the current state of the network in terms of both loss probability and delay. To the best of our knowledge, this is the first work which deals with the route selection for deflected bursts using a learning-based approach; (2) a novel multi-path routing (proactive) approach, called Reinforcement Learning-based Multi-path Routing (RLMR), to reduce loss probability before wavelength contentions happen. RLMR is a distributed scheme; it is concerned with the selection of an optimal path in terms of both burst loss probability and end-to-end delay. Each node in the network has to learn the best output link to route an incoming burst for a given destination based on the current state of the network in terms of both loss probability and delay. The operation of RLMR establishes a suitable load balancing in OBS networks; and (3) a (proactive and reactive) scheme, called Integrated Reinforcement Learning-based Routing and Contention Resolution (IRLRCR), to reduce loss probability by integrating RLDRS and RLMR. The three schemes are distributed, scalable and well adapted to dynamic changes in the network state and topology.

The remainder of this paper is organized as follows. Section 3.2 presents related work on multi-path routing and deflection routing for OBS networks. Section 3.3 provides a short overview of Reinforcement Learning and Q-learning. Section 3.4 describes the proposed deflection routing scheme called RLDRS. Section 3.5 describes the proposed multi-path routing approach called RLMR. In Section 3.6, we describe IRLRCR that integrates RLMR and RLDRS. In Section 3.7, we present simulation results that show the performance of RLMR, RLDRS and IRLRCR. Finally, Section 3.8 concludes the paper.

3.2. Related work

In the literature of OBS networks, we find a number of contributions which have studied multi-path routing [36, 37, 39, 40, 85]. In these contributions, a set of paths is a priori known between each source and destination pair of nodes. In this case, the choice of the routing path to a given destination is performed at the source node. Hence, the problem is what information to use, how to update it and how to determine the optimal path given this information. The authors in [85] consider several metrics (e.g. link utilization, path end-to-end delay or burst loss probability) to evaluate the level of congestion in a given path. The choice of the optimal path could be based on pure strategy (if a single metric is considered) or hybrid (if two or more metrics are combined). The authors conclude that, in general, path switching outperforms shortest path routing that is usually used in OBS networks. In [39], paths from the same source to the same destination are assigned priorities; bursts are sent using a path that has the highest priority. Each time a data burst is sent on a given path, the source node receives a feedback which indicates either the success or the failure of the burst transmission so that it can update path priorities. To accelerate the rate of priorities updating, each time a data burst is sent, search packets are sent on the other paths leading to the same destination (using control wavelengths) in order to receive feedbacks indicating whether transmitting the data burst on the other paths would be successful or not. We note that search packets may cause a considerable control overhead. The authors in [37] propose a congestion avoidance-based scheme in which each network core node measures the load on each one of its output links and sends, periodically, this information to all of the network edge nodes; these nodes use the load information to avoid congested paths. The authors propose another scheme that considers a weighted function based on congestion and hop count (between source and destination) to select paths that improve performance in terms of burst losses and delay. In [40], a passive probing on sub-optimal paths is used; the source node sends a small fraction of its traffic on these paths while keeping the probabilities of choosing them very low and analyzes feedback ACK and NACK packets during a time sliding window to determine an optimal path. A burst pipelining scheme is also proposed to guarantee in-order burst delivery. More recently, the authors in [36] proposed to put reinforcement learning agents using Q-learning at each ingress edge node. An agent chooses for every burst, to be transmitted, the optimal path in terms of burst loss probability and use ACK and NACK feedback packets to update its appreciation of paths. In this paper, RLMR uses learning agents at both edge and core nodes to choose optimal links rather than an optimal path; this results in a finer granularity path selection. In addition, RLMR considers both burst loss probability and end-to-end delay as metrics for the path selection process. Furthermore, control overhead is considerably reduced since RLMR uses only one-hop feedback packets of very small size.

Deflection routing has been considered since the early time of OBS networks as a cost-effective yet efficient contention resolution approach. However, early studies focused on regular topologies and traffic with particular characteristics [47]. In [48], the authors propose a deflection routing protocol that uses two functions: sender check and sender retransmission which aim to control deflection. In [49, 88], the authors study the performance of Just Enough Time (JET)-based OBS networks with deflection routing and FDLs; they reported that deflection routing brings a performance gain with fewer wavelengths and lightly loaded network. In [50] the authors study burst loss performance of an OBS node with deflection routing and with and without wavelength conversion. According to the authors, unlike other models which suppose a single deflection output port (e.g., [49]), this model can be applied to an OBS node with any number of output ports. In [51], a wavelength reservation approach for deflection routing is proposed. This approach is analogous to trunk reservation in circuit switched networks and intentionally limits the amount of deflections at high loads by reserving a number of wavelengths for the exclusive use of non-deflected bursts. In [52], Shortest Path Prioritized Random Deflection Routing (SP-PRDR) is proposed where deflection to a random neighbor is performed in case of contention. The deflected burst is assigned a low priority contrary to non-deflected bursts which have a high priority by default. Hence, in case of contention between a deflected burst and a non-deflected burst, the later is privileged and the deflected burst is dropped if necessary. The authors in [53] propose a scheme called Contention-based Limited Deflection Routing (CLDR) where two functions are performed: (a) determining whether a burst should be deflected or retransmitted from the source based on some performance criteria; and (b) If the decision is to deflect a burst, then an optimal path that minimizes both distance and loss due to contention is selected. The second function (b) is based on an optimization model that has to be resolved periodically. In [54], the authors propose an optimization model to dynamically determine the deflection path with the least traffic load.

Moreover, bursts with high priority are assigned wavelengths virtually over the deflection path. The authors in [55] proposed three deflection routing strategies, called Least Loaded Node (LLN), LLN/U-turn and LLN/U-turn/ LB. LLN deflects a burst to the least loaded neighbor except the neighbor from which it has arrived. U-turn option allows deflecting a burst to the neighbor from which it has arrived and LLN/U-turn/LB uses LLN/U-turn with a load balancing scheme [56]. Authors reported that U-turn and LB options bring a substantial reduction of burst loss probability without a significant increase of mean burst end-to-end delay. In [57], a deflection routing scheme that considers the possibility of contention at downstream nodes is proposed. A metric called "expected route distance" towards destination node is defined. This metric considers the possibility of contention at each downstream transit node. It is calculated using measured link loss probabilities at each downstream transit node and used to select a deflection output link. From deflection output link perspective, we can classify existing deflection schemes in three categories: (a) static: deflection paths are computed using static parameters, such as in [49, 51, 88]; (b) random: deflection paths are computed randomly, such as in [48, 50, 52]; and (c) dynamic: deflection paths are computed using dynamic parameters (e.g., QoS metrics and link load) such as in [53-55, 57]. In this paper, we propose a dynamic scheme that selects an optimal deflection link toward a given destination using reinforcement learning techniques; the objective is to outperform existing schemes, in terms of burst losses caused by contention, with less overhead.

3.3. Reinforcement learning

In the framework of reinforcement learning, a learning agent interacts with its environment by accepting inputs from it (e.g., loss probability on an output link) and responding by selecting appropriate actions (e.g., selecting the best output link toward a destination to send a burst). The environment evaluates the agent decisions by sending a rewarding or penalizing reinforcement feedback. Based on the value of the received reinforcement feedback (also called *signal*), the agent updates its parameters so that good decisions become more likely to be made in the future, while bad decisions become less likely to occur [89]. In addition to the learning agent and the environment, the key components of a reinforcement learning system are: (a) A policy which defines the behavior of the learning agent at a given time; (b) A reward function that maps each state or state-action pair of the environment to a numerical value which indicates the desirability of that state or the desirability of an action at that state; (c) A value function which represents the expected cumulative reward in the future starting from a given state (the objective of the learning agent is to maximize this cumulative reward over the long run); and (d) A model of the behavior of the environment which can predict the new state and the reward when the agent performs a given action; this component is optional.

Q-learning is a well-known algorithm of reinforcement learning framework [90]. A Q-learning agent learns the best action to take in a given state by trying repeatedly all possible actions in that state and by evaluating the estimated long-term discounted reward [91]. This reward is called *Q-value* noted $Q^{\pi}(s,a)$ which represents an estimate of choosing action a at state s and following policy π thereafter. With Q-learning, the objective is to find the optimal policy noted π^* by estimating its Q-value. The core of Q-learning algorithm is given in (3.1):

$$Q^{\pi}(s_t,a_t) \leftarrow Q^{\pi}(s_t,a_t) + \alpha(r_{t+1} + \gamma \max_{a} Q^{\pi}(s_{t+1},a) - Q^{\pi}(s_t,a_t)) \tag{3.1}$$

where α is the learning rate, γ is the discount factor, s_t, a_t and r_t are the state, the action and the reward at time t, respectively (see [90] for more details).

3.4. The proposed deflection routing scheme

In this section, we present the details of our deflection routing scheme (RLDRS) [87]. First, we present the Q-learning based scheme that will be used at each OBS network node to determine the optimal alternative output link to deflect an incoming burst when needed. Then, we discuss issues related to reinforcement learning, namely, exploration versus exploitation and convergence. After that, we present overhead analysis of the proposed scheme. Finally, we discuss Insufficient Offset Time Problem (IOT).

The objective of our proposal is to find an optimal output link to deflect a burst when only deflection routing is used as a contention resolution approach. To the best of our knowledge, our scheme is the first to use reinforcement learning to deal with the deflection path selection problem. With our scheme, each node in the network learns optimal deflection output links with far less (communication and computation) overhead than existing schemes.

3.4.1. The proposed scheme

We propose to make the selection of deflection output link, toward a given destination, by each node in the network, adaptive based on the current state of the network. We suppose that each node has a learning agent that learns, continuously, an optimal deflection output link towards a given destination at a given time. A learning agent uses a lookup table called Deflection Table (DT) to store values (called Q-values) representing its appreciation of a deflection output link with respect to a destination. This appreciation takes into consideration both burst loss probability and delay (in terms of hops) experienced by bursts from the current node toward the destination through this output link. Thus, each entry in the node's Deflection Table is indexed by the pair (destination, neighbor); the computation of Q-value will be described later in the section. When a wavelength contention occurs in a node and a burst has to be deflected, the learning agent of this node decides to forward the burst to a neighbor with the highest Q-value (other than the node of the primary path). For example, if a node x decides to deflect a burst with destination d, it forwards the burst to its neighbor y determined in (3.2):

$$y = \underset{z \in N(x)\, and\, z \neq p}{\operatorname{argmax}} Q_x(d, z) \tag{3.2}$$

where argmax stands for argument of the maximum (namely, the neighbor with the maximum Q-value, other than the neighbor on the primary route noted p, with respect to a given destination), $Q_x(sd,z)$ is the Q-value associated with neighbor z and destination d, and $N(x)$ is the set of neighbors of node x.

Initially, we assume that the loss probability of each output link in the OBS network is null. To initialize the deflection table of a node, say x, each Q-value corresponding to a given destination, say d, and a given neighbor, say y, is computed based on the shortest path delay (the number of hops) between x and d where the first hop neighbor is y. Thus, we ensure that if the loss probability is very low or negligible, RLDRS converges to Shortest Path Deflection Routing (SPDR) (i.e. deflect bursts to the second shortest path route to destination).

Whenever a node deflects a control packet to a neighbor, it receives a feedback packet from the neighbor that it uses to update the neighbor's (and corresponding destination) entry in the deflection table. For example, in Figure 9, when node x deflects a control packet with destination d to its neighbor y (rather than sending it to the primary node i), it receives a feedback packet from y which contains a numerical value f_{yx} defined in (3.3):

$$f_{yx} = Q_y(d,z) \cdot D_y(d,z) \tag{3.3}$$

where node z is chosen and $Q_y(d,z)$ is computed by node y using (3.2), and $D_y(d,z)$ is the delay between node y and node d through neighbor z (the delay is the number of hops of the shortest path between node z and node d). We assume that each node in the OBS network knows the number of hops of the shortest path between itself and each destination through each of its neighbors. The multiplication of $Q_y(d,z)$ by $D_y(d,z)$ to calculate the feedback value f_{yx} is necessary to eliminate the delay factor already considered in $Q_y(d,z)$;

this delay is not useful for the destination node of the feedback packet (node x in this case), that will use its own delay to the destination node d (see (3.4)).

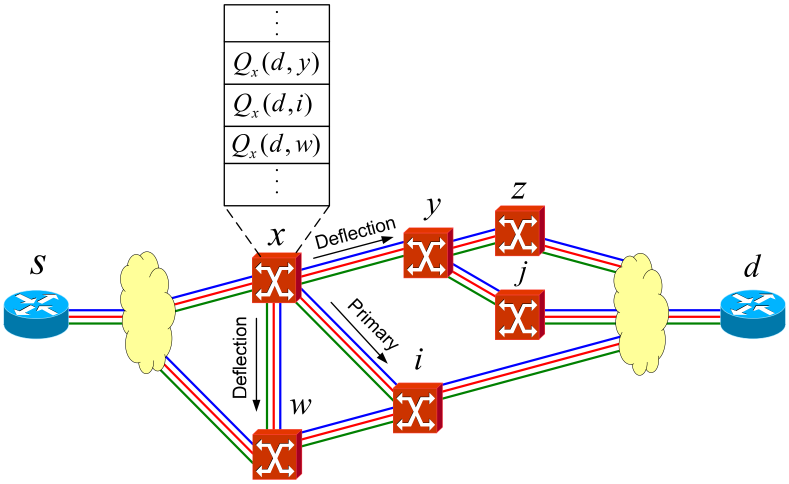


Figure 9. An Example that shows a part of node x's deflection table.

Upon receipt of the feedback packet, node x updates its deflection table as shown in (3.4):

$$Q_x(d,y) \leftarrow Q_x(d,y) + \alpha((f_{yx}(1-B_{xy})/D_x(d,y)) - Q_x(d,y)) \tag{3.4}$$

where $0 < \alpha \leq 1$ is the learning rate, B_{xy} is the burst loss probability on the output link from node x to node y. It is measured using a time sliding window; at the end of each time window of duration τ, B_{xy} is calculated as shown in (3.5):

$$B_{xy} = \begin{cases} \dfrac{Drop_{xy}}{Sent_{xy} + Drop_{xy}}; & \text{If } Sent_{xy} + Drop_{xy} > 0 \\ 0; & \text{If } Sent_{xy} + Drop_{xy} = 0 \end{cases} \tag{3.5}$$

where $Drop_{xy}$ and $Sent_{xy}$ are the number of dropped bursts and successfully transmitted bursts through the output link from node x to node y during the last time window, respectively.

The idea behind (3.4) is to estimate the probability that a burst will be dropped along the path from x to d through y. Indeed, assuming that link drop probabilities are independent, the probability that a burst will be dropped through a path p consisting of nodes $n_1,...,n_{|p|}$ (noted b_p), is calculated using (3.6):

$$b_p = 1 - \prod_{1 \le i \le |P|-1} (1 - B_{n_i n_{i+1}}) \tag{3.6}$$

In (3.4), we introduce the delay (number of hops) of the path, under consideration, to make a tradeoff between loss probability and delay. Thus, shorter paths are preferred over longer paths when the longer path's improvement in terms of loss probability is not substantial.

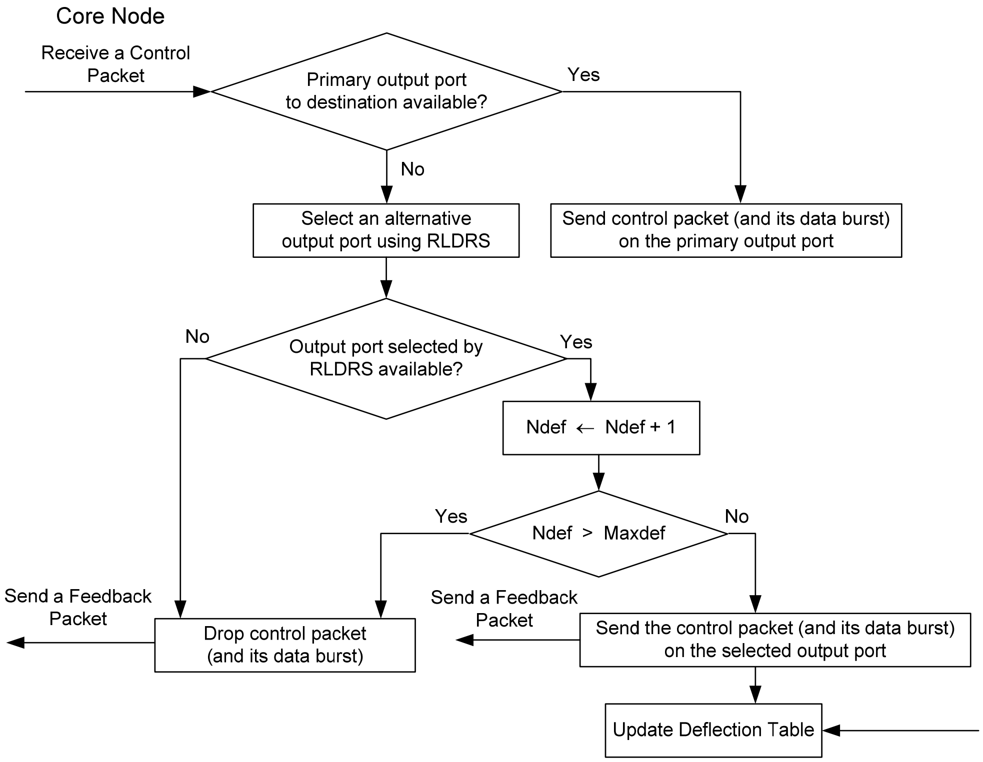


Figure 10. The operation of RLDRS at an OBS core node.

To control the additional traffic in the network caused by deflection routing and prohibit excessive deflections, we fix the value of the maximum number of authorized deflections to *Maxdef* and we add a new field, noted *Ndef*, to the control packet. Each time the burst is deflected, the value of *Ndef* is incremented. When *Ndef* reaches *Maxdef* the burst is simply dropped. Figure 10 shows the operation of RLDRS at an OBS core node.

3.4.2. Exploration, convergence and overhead analysis

RLDRS aims to exploit what a learning agent has learnt before (from its interaction with the environment) by selecting the neighbor with the highest Q-value in the deflection table, to forward an incoming burst; this is called exploitation. Nevertheless, an exploration policy is required to check whether another neighbor becomes better than the current optimal deflection neighbor due to changes in the network state (e.g., traffic pattern, level of contentions, network topology, etc.). Exploration policy is, also, of high importance at the beginning of the network operation, where deflection tables are initialized based on Shortest Path Routing and learning agents try to find the optimal deflection neighbors based on the current network state. Hence, we adopt an $\varepsilon - greedy$ exploration policy [91] which makes the decision process (the selection of a neighbor to deflect an incoming burst) probabilistic, with a small probability ε to select a non optimal deflection neighbor (e. g., $\varepsilon = 0.1$), and a high probability $(1-\varepsilon)$ to select an optimal deflection neighbor (e.g., $\varepsilon = 0.9$).

Convergence is a known issue in reinforcement learning. Indeed, when a reinforcement learning model is applied to solve a problem, it is not always guaranteed that this model will converge to a stable solution. Fortunately, according to [90], a Q-learning algorithm that uses a lookup table representation of estimates is guaranteed to converge, which is the case for RLDRS.

To update deflection tables, each time a node deflects a burst to a neighbor, a feedback packet is sent back to the node by the neighbor. We note that the overhead caused by feedback packets is negligible and has no effect on the performance of OBS network since: (a) a feedback packet is sent only when a burst is deflected; (b) feedback packets are sent on wavelength(s) used only to transmit control traffic; and (c) the size of feedbacks is very small (a numerical value).

3.4.3. Insufficient offset time problem

An issue related to buffer-less OBS networks when deflection routing is adopted is known as Insufficient Offset Time (IOT). This problem occurs when the length of routing paths is not known by edge nodes. In this case, the Offset Time (OT) that has to separate the control packet from its data burst may be underestimated which could result in the drop of the data burst after it surpasses its control packet, before reaching the destination. This problem may occur in RLDRS since deflected bursts will, probably, take paths longer than the primary path which is the shortest path.

The simplest solution to this problem is to add a fixed delay to the OT. The value of this delay can be estimated as the delay difference between the shortest path and the longest path between the source node and the destination node. We note that the determination of the longest path requires that the deflection scheme be loop-free and that the authorized number of deflections be limited. It is worth noting that the added fixed delay will not have a considerable impact on the average end-to-end delay in the OBS network; indeed, the impact consists of increasing the end-to-end delay by the control packet processing time multiplied by the number of additional hops if any; the control packet processing time is very small compared to the mean burst assembly time, for example. A more sophisticated approach can be proposed to determine adaptively the value of this additional delay by estimating the current length of the path between two nodes in the network at the cost of additional control overhead; however, this is out of the scope of this paper.

3.5. The proposed multi-path routing approach

In this section, we present the details of a novel distributed multi-path routing approach, called Reinforcement Learning-based Multi-path Routing (RLMR).

RLMR is based on a reinforcement learning scheme similar to the one used in RLDRS. However, RLMR is different from RLDRS in that it can route data bursts on anyone of the links at each node between the source and the destination rather than on the shortest path as in RLDRS. Moreover, in case of wavelength contention, RLMR simply drops one of the two bursts involved in the contention rather than deflecting it to an alternative available output link as in RLDRS. First, we present the operation of RLMR. Then, we discuss exploration and convergence issues and present overhead analysis of the proposed scheme. Finally, we present our loopless forwarding algorithm and discuss out-of-order delivery and Insufficient Offset Time (IOT) problems.

3.5.1. The routing scheme

The deflection routing scheme of RLDRS can be used to route each burst from source to destination (without considering contention resolution). In RLMR, the path selection process is distributed on all the nodes in the network, i.e., any node in the network has to decide which link a burst should take to reach its destination. Moreover, each node in the network has a learning agent that learns the optimal output link to forward a burst to a given destination at a given time. A learning agent uses a lookup table called *Q-table* to store values (called Q-values) representing its appreciation of an output link with respect to a destination. This appreciation takes into consideration both burst loss probability and delay (in terms of hops) experienced by bursts from the current node toward the destination through a chosen output link. Hence, each entry in the node's Q-table is indexed by the pair *(destination, neighbor)*; the computation of Q-value is similar to RLDRS and will be described later in the section. When a node receives an incoming burst, the learning agent of this node decides to forward this burst to a neighbor with the highest Q-value. For example, if a node x receives a data burst with destination d, it forwards that burst to its neighbor y determined in (3.2).

Initially, we assume that the loss probability of each output link in the OBS network is null. To initialize the Q-table of a node, say x, each Q-value corresponding to a given destination, say d, and a given neighbor, say y, is computed based on the shortest path delay (the number of hops) between x and d where the first hop neighbor is y. Thus, we ensure that if loss probability is very low or negligible, RLMR converges to *Shortest Path Routing*.

Whenever a node sends/forwards a control packet to a neighbor, it receives a feedback packet, from that neighbor, that it uses to update the neighbor's (and corresponding destination) entry in the Q-table. For example, in Figure 11, when node x sends a control packet with destination d to its neighbor y, it receives a feedback packet from y which contains a numerical value f_{yx} defined in (3.3). We assume that each node in the OBS network, knows the number of hops of the shortest path between it and each destination through each of its neighbors.

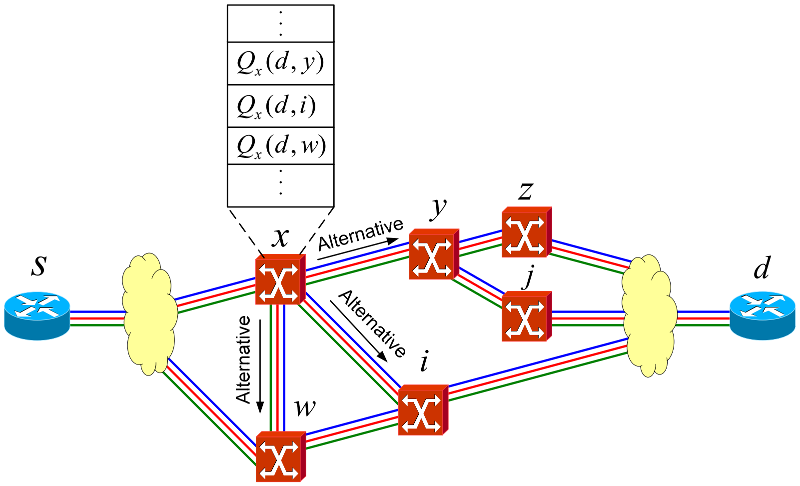


Figure 11. An example that shows a part of node x's Q-table.

Upon receipt of the feedback packet, node x updates its Q-table (Q-value) as shown in (3.4). Figure 12 shows the operation of RLMR at an OBS core node.

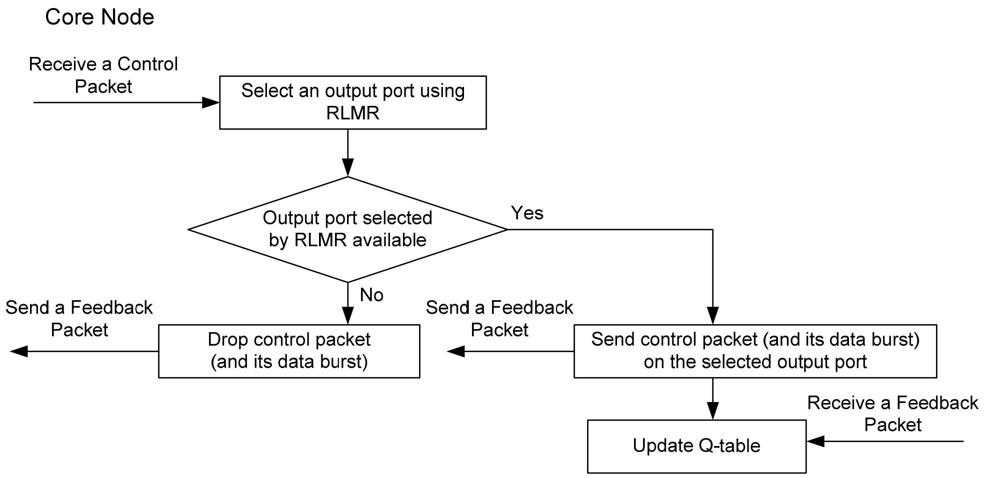


Figure 12. The operation of RLMR at an OBS core node.

3.5.2. Exploration, convergence and overhead analysis

In RLMR, we adopt an $\varepsilon - greedy$ exploration policy as explained for RLDRS. Moreover, the convergence of RLMR is guaranteed since it uses a lookup table representation of estimates, which is the Q-table.

To update Q-tables, each time a node forwards a burst to a neighbor, a feedback packet is sent back to the node by the neighbor; this may seem causing considerable overhead. Fortunately, this overhead has almost no effect on the performance of the OBS network since: (a) feedback packets are sent using wavelength(s) used only to transmit control traffic; and (b) the size of feedbacks is very small (a numerical value). Furthermore, no routing is needed since all feedback packets traverse no more than one hop. This is different from existing approaches where an ACK packet is returned back to the source node if the burst is successfully received by the destination node, and a NACK packet is returned back to source node if the burst is dropped by an intermediate node (e.g., [36, 39, 85]). These ACK/NACK packets are considerably larger (than control packets in RLMR) since they include routing information from the sending node (destination or an intermediate node) to the source. Thus, the overhead introduced by RLMR is considerably less important than existing approaches.

3.5.3. Loopless routing, out-of-order delivery and insufficient offset time

When distributed routing is used, loops may occur in routing paths, i.e., a packet (burst) may pass through the same node more than once. In the worst case, a packet may circulate indefinitely in a loop causing an increase in the network load and, consequently, an increase in the burst loss probability. In order to prevent loops from appearing in RLMR, we formulate the problem of computing a set of alternatives to forward incoming bursts in each node (with respect to their destinations) as a graph theory problem. We propose to modify the algorithm reported in [92] (which aims to find one deflection alternative in each node with respect to a given destination) to construct as many alternatives as possible to forward bursts towards a given destination. We consider the OBS network as a graph $G(V,E)$ where V is the set of nodes and E is the set of directed links. We associate to each node v a directed spanning tree T_v, rooted at v and having edges E_{T_v} directed towards v. T_v is obtained by selecting for each node (other than v) in the network its outgoing link that is traversed by the shortest path between this node and v. This link can be obtained by running Dijkstra's algorithm between each two nodes in the network and then by picking up the first hop link. Having the graph $G(V,E)$ and the spanning tree T_v for each node v, the problem consists of determining a set of links $E_{F_v} \subset E$ such that the directed graph $F_v = (V, E_{T_v} \cup E_{F_v})$ is acyclic. Algorithm 1 shows the pseudo-code of the proposed loopless forwarding algorithm.

The algorithm selects, each time, only one node and constructs a spanning tree that will be rooted at the selected node. The role of the spanning tree is, on one hand, to ensure that the link on the shortest path between each other node and the selected node will be in the list of forwarding alternatives. On the other hand, the spanning tree creates a topological sorting between network nodes with respect to the selected node. Then, in step 2, the node with the minimal nodal degree is selected from the farthest level in the spanning tree in order to alleviate the case when one node has many forwarding alternatives and the other nodes have no forwarding alternatives. In step 3, all the alternatives of the selected node are added to the list of forwarding alternatives. In step 4, we remove all incoming links to the selected node from the list of forwarding alternatives; this ensures that no burst forwarded from the selected node will return to it in the future.

The process described above will continue for each node in each level (of the spanning tree) until all nodes are exhausted. In addition to guaranteeing a loop-free routing, this algorithm can be used with delay sensitive traffic to ensure that bursts of this class of traffic will not experience an excessive additional delay in the core of the OBS network.

Out-of-order delivery is a well-known issue in multi-path and distributed routing. In [40], the authors propose a solution that consists of implementing a pipelining process where a delay is added between the transmissions of two successive bursts sent on different paths. They show that their pipelining process does not decrease the throughput in the network. We can adopt this pipelining process to our distributed routing by adding a fixed additional delay between the transmissions of two successive bursts.

The problem of *insufficient offset time* may occur in RLMR when the length of routing paths is longer than the estimated length, made by the edge node. In RLMR, we can tackle this problem using the same approaches proposed for RLDRS, namely: (a) add a fixed delay to the basic Offset Time (without a considerable impact on the average end-to-end delay, as mentioned in Section IV); and (b) use a more sophisticated approach to determine adaptively the value of this additional delay by estimating the current length of the path between two nodes in the network. This second approach is out of the scope of this paper.

Algorithm 1. Loopless forwarding algorithm.

```
Input:
    A graph $G(V,E)$;
    $|V|$ spanning trees $T_v = (V, E_{T_v})$ each rooted at unique node $v \in V$;
Output:
    $E_{F_v}$, $\forall v \in V$, the set of forwarding edges (links) for node $v$;
    $F_v = (V, E_{T_v} \cup E_{F_v})$, $\forall v \in V$, a directed forwarding graph;
Step 0:
    Set $V^* = V$;
    Set $E_{F_v} = \{\ \}$, $\forall v \in V$;
Step 1:
    Select a vertex $v \in V^*$;
    Set $E^* = \{E - E_{T_v}\}$;
    Let $d_v$ be the depth of tree $T_v$;
    Let $L_i$, $i \in \{1, 2, ..., d_v\}$ be the set of nodes which are at distance $i$ from the root node $v$;
    Set $k = d_v$;
Step 2:
    Select node $u \in L_k$ such that the degree of node $u$, $\delta(u) = \min_{u' \in L_K} \delta(u')$;
Step 3:
    Set $E_{F_v} = E_{F_v} \cup e(u,w)$, $\forall e(u,w) \in E^*$ directed link between nodes $u$ and $w$;
Step 4:
    Remove all links $e(w,u)$, $\forall w \in V$ from $E^*$;
    Remove node $u$ from $L_k$;
    If $L_k = \{\ \}$ then $k = k - 1$;
    If $k \neq 0$ then go to step 2, otherwise, go to step 5;
Step 5:
    Set $F_v = (V, E_{T_v} \cup E_{F_v})$;
    Remove node $v$ from $V^*$;
    If $V^* \neq \{\ \}$ then go to step 1, otherwise STOP;
```

3.6. The integrated approach

The integrated approach, called *Integrated Reinforcement Learning-based Routing and Contention Resolution* (IRLRCR) approach, aims to adopt both a proactive approach and a reactive approach to reduce burst losses. In IRLRCR, RLMR is adopted as the proactive approach and RLDRS is adopted as the reactive approach. This combination works well since RLMR and RLDRS adopt the same general approach adopted by Q-learning algorithm with lookup table representation of estimates to take decisions. Indeed, in IRLRCR, we use only one lookup table, called *Global Table*, which is used by both RLMR and RLDRS as the Q-table and the Deflection Table, respectively. In the normal case, bursts are routed using RLMR; however, when a contention occurs and a burst has to be deflected, RLDRS is used. Thus, since both RLMR and RLDRS use the same lookup table, the neighbor with the second highest Q-value is selected by RLDRS to deflect a burst.

Figure 13 shows the operation of IRLRCR in an OBS core node. Whenever a burst is received by an OBS node, RLMR is used to determine the output port towards its destination. If this output port is not available (i.e. there is a wavelength contention), RLDRS is used to determine an alternative (deflection) output port. If this output port is not available too, the incoming burst is simply dropped. Otherwise, the burst is sent on the selected output port. Nevertheless, because of the $\varepsilon - greedy$ exploration policy used by both RLMR and RLDRS, even if it is very small, the possibility that RLMR and RLDRS select the same output port in the same selection process exists. To tackle this problem, we add a simple test to check whether the output port selected by RLMR is the same as the one selected by RLDRS. If it is the case, another output port is selected. The new output port is the one with the highest Q-value other than the current output port (selected by both RLMR and RLDRS).

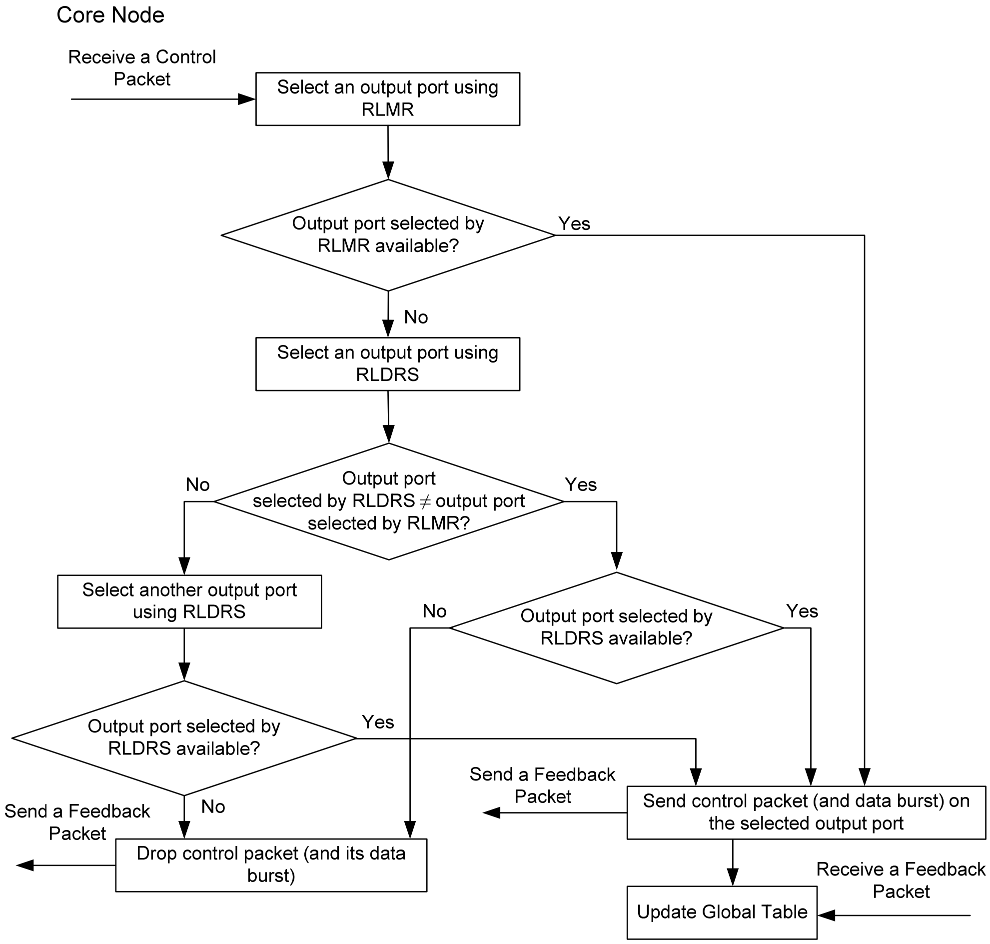
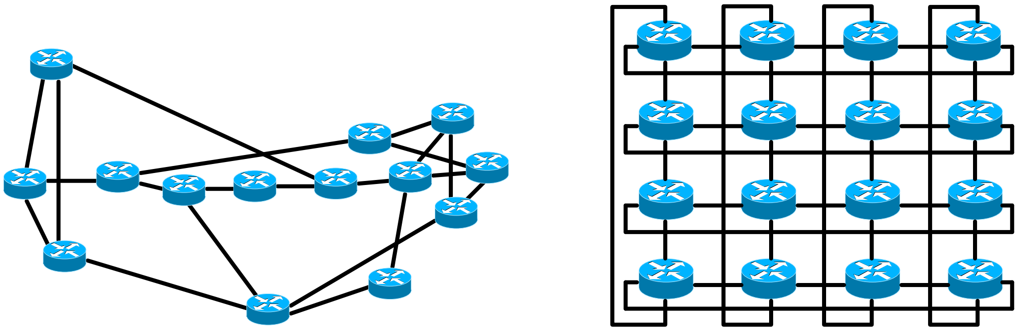


Figure 13. The operation of IRLRCR at an OBS core node.

3.7. Simulation results and analysis

In this section, we present simulation results that we have performed to evaluate the performance of RLDRS, RLMR and IRLRCR. We use the ns-2 simulator [93] and modules that implement OBS in ns-2 [94]. We consider two kinds of topologies, namely, mesh topologies represented by NSFNET with 14 nodes and regular topologies represented by regular 4 x 4 nodes torus topology (Figure 14). We assume that each single fiber link is bidirectional and has the same number of wavelengths. Each node in the network can generate, route and receive traffic. Sources and destinations of traffic connections are generated randomly between any two nodes in the network. The traffic load is expressed as the average traffic load per link. The capacity of a link is the sum of the capacities of all the wavelengths in this link. We use *Min Burst length Max Assembly Period* (MBMAP) algorithm for burst assembly [30], with maximum burst size fixed to 10 KB (Kilo Bytes) and LAUC-VF (Last Available Unused Channel with Void Filling) algorithm for wavelength assignment in OBS edge nodes. In our simulations, we use exponential ON/OFF traffic.



(a) (b)

Figure 14. NSFNET topology with 14 nodes (a) and regular torus topology with 4 x 4 nodes (b).

Initially we set (a) the learning rate α to 0.6 (other values of α have a slight impact on the trend of the simulation results; they will not be presented here); (b) the period to measure the loss rate on each link (B_{xy}) to 2 s; and (c) the exploration probability ε to 0.02. For QPS parameters, we use the same values reported in [36]. All the following results have a confidence level of 95%.

The goal of these simulations is to measure the performance of RLDRS, RLMR and IRLRCR in terms of loss probability which is the main performance metric in buffer-less OBS networks. In addition, since RLDRS, RLMR and IRLRCR use longer routing paths than *Shortest Path Routing*, we evaluate them in terms of mean burst end-to-end delay. In this paper, the end-to-end delay represents the hop distance between the source node and the destination node. We make a reasonable hypothesis that the network is properly designed to have proportionate link propagation delays. This means that we do not have very long links and very short links in the same network.

We compare RLDRS to both (a) *Shortest Path First (SPF)* routing which represents the original buffer-less OBS with shortest path routing and without deflection routing and (b) *Shortest Path Deflection Routing (SPDR)* where the selected deflection output link is the first hop on the second shortest path towards the destination node. It is worth noting that several of the proposed deflection routing schemes have proposed or used SPDR as a reference scheme [49, 51, 54, 55, 88]. Also, we set $Maxdef$ to 2 (that means that a burst is not authorized to be deflected more than two times).

We compare RLMR to both *Shortest Path First (SPF)* routing and *Q-learning algorithm for Path Selection (QPS)* proposed in [36]; the motivations behind this choice are (a) SPF is considered as the standard routing algorithm, not only in OBS networks, but also in most of data communication networks; (b) QPS is the most recent work on multi-path routing in OBS networks (to the best of our knowledge) that outperforms existing approaches (e.g., [85]); and (c) QPS represents the first attempt to use reinforcement learning in the routing of OBS networks [36].

We compare the performance of IRLRCR to the performance of RLMR and the performance of RLDRS. We do not use the loopless forwarding algorithm with IRLRCR; we found that this algorithm decreases the contribution of RLDRS to improve IRLRCR performance. Overall, we found that when loopless forwarding is used, the performance of IRLRCR is comparable to the performance of RLDRS in terms of loss probability and comparable to the performance of RLMR in terms of end-to-end delay. Nevertheless, for traffic flows with firm end-to-end delay requirement, loopless forwarding algorithm can be used to guarantee an upper bound on end-to-end delay at the cost of a possible increase in burst loss probability.

Figure 15 shows burst loss probability when varying the load on NSFNET. We can see clearly that RLDRS outperforms both SPF and SPDR. Indeed, The relative improvement (defined as [(Loss with SPF (SPDR) – Loss with RLDRS) / (Loss with SPF (SPDR))]) of RLDRS compared to SPF is about 94% at load 10% and about 65% at load 100% and compared to SPDR it is about 91% at load 10% and about 59% at load 100%. Overall, the average relative improvement of RLDRS over SPF is about 77% and over SPDR is about 72%. This proves that RLDRS considerably decreases loss probability of OBS network with deflection routing.

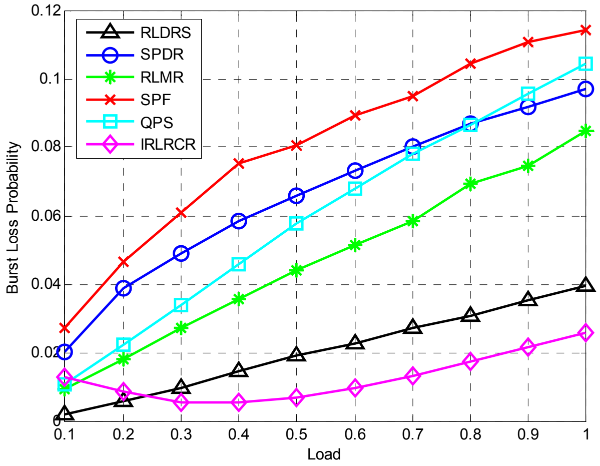


Figure 15. Loss probability vs. load on NSFNET with 64 wavelengths.

For RLMR, Figure 15 shows clearly that RLMR outperforms both QPS and SPF. Indeed, The relative improvement (defined as [(Loss with SPF (QPS) – Loss with RLMR) / (Loss with SPF (QPS))]) of RLMR compared to SPF is about 66% at load 10% and about 25% at load 100% and compared to QPS it is about 15% at load 10% and about 18% at load 100%. Also, we can see that, overall, IRLRCR clearly outperforms RLMR and RLDRS. Indeed, at load 100%, the relative improvements of IRLRCR over RLMR and RLDRS are about 69% and 35%, respectively.

Figure 16 shows that SPF has, generally, the lowest mean end-to-end delay, which is expected. SPDR has similar performance to SPF and it outperforms RLDRS. This is expected since SPDR deflects bursts based, exclusively, on shortest paths to destination. While the mean burst end-to-end delay (over all of the loads) is around 2.1 for SPF and SPDR, it is around 2.25 for RLDRS. This additional delay is acceptable if we consider the significant improvement of RLDRS in terms of loss probability.

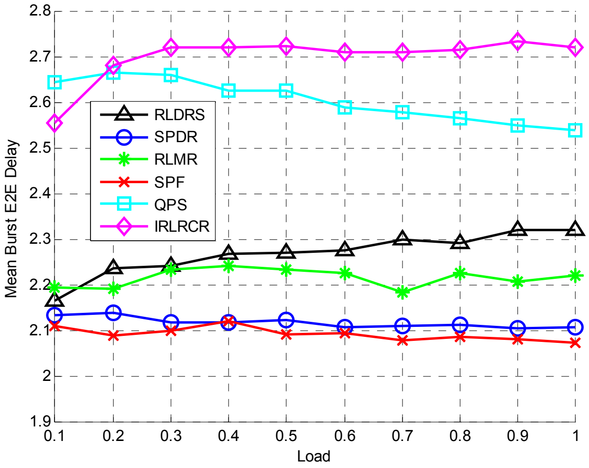


Figure 16. Mean burst end-to-end delay (number of hops) vs. load on NSFNET with 64 wavelengths.

Figure 16 also shows that RLMR improves significantly end-to-end delay compared to QPS. Indeed, whereas mean burst end-to-end delay is around 2.6 for QPS, it is around 2.2 for RLMR. Moreover, we can see that IRLRCR has the highest mean end-to-end delay. Indeed, the mean end-to-end delay of IRLRCR is around 2.7. This additional delay of IRLRCR is expected, since it uses both RLMR and RLDRS which add the additional delay of each of them to the delay of IRLRCR. However, this delay is acceptable for medium and small size networks and for traffic flows with no firm end-to-end delay requirement, especially, if we consider the considerable improvement of IRLRCR over both RLMR and RLDRS in terms of loss probability.

Figure 17 shows the loss probability when varying the load in the regular 4 x 4 topology. At load 100% the relative improvement of RLDRS when compared to SPF is about 59% and compared to SPDR is about 41%. The average relative improvement of

RLDRS compared to SPF is about 73% and compared to SPDR is about 60%. This proves that regardless of the type of topology (mesh or regular), RLDRS outperforms SPF and SPDR by considerably reducing loss probability.

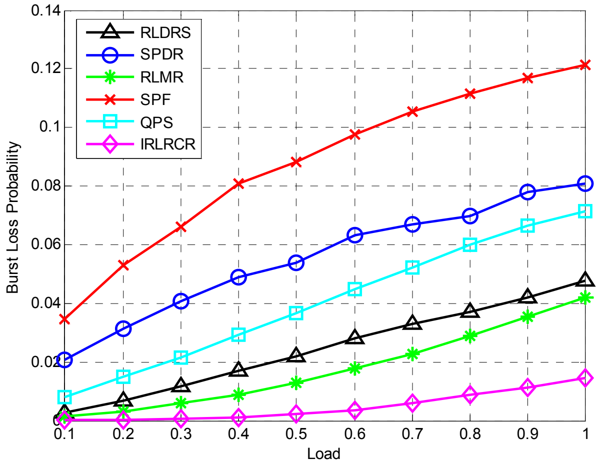


Figure 17. Loss probability vs. load on regular 4 x 4 topology with 64 wavelengths.

Also, we can see that the improvement of RLMR is better using this topology compared to NSFNET; at load 100% the relative improvement of RLMR when compared to SPF is 65% and compared to QPS is 41%. This can be explained by the fact that in this topology, the average node degree is 4, whereas the average node degree in NSFNET is 3. This supports the idea that RLMR performs better whenever the average node degree increases, due to the increase in the number of forwarding alternatives in each node in the OBS network. Figure 17 also shows that IRLRCR clearly outperforms both RLMR and RLDRS (with mean relative improvements 80% and 86%, respectively). Besides, we observe that RLMR outperforms RLDRS. We explain that by the fact that RLMR performs better when the mean node degree increases. This confirms our findings that whenever the mean node degree increases, multi-path routing becomes more and more efficient in reducing loss probability.

Figure 18 shows mean burst end-to-end delay with regular 4 x 4 topology. Here again, and for the same reasons, we observe the same behaviors as in Figure 16, namely, RLDRS underperforms SPF and SPDR, RLMR outperforms QPS and slightly underperforms SPF and IRLRCR underperforms RLMR and RLDRS.

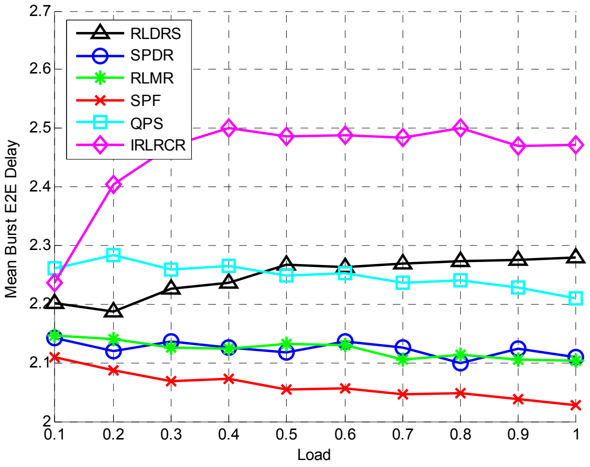


Figure 18. Mean burst end-to-end delay (number of hops) vs. load on regular 4 x 4 topology with 64 wavelengths.

Figure 19 shows burst loss probability when fixing the traffic load at 60% and varying the number of wavelengths from 8 to 128 wavelengths on NSFNET. We can see clearly that regardless of the number of wavelengths: (a) RLDRS outperforms both SPF and SPDR; (b) RLMR outperforms SPF and QPS; and (c) IRLRCR outperforms both RLMR and RLDRS.

Figure 20 shows burst end-to-end delay when fixing the traffic load at 60% and varying the number of wavelengths from 8 to 128 wavelengths on NSFNET. Here again, we see that: RLDRS, RLMR and IRLRCR reduces effectively loss probability at the cost of slight increase in end-to-end delay regardless of the number of wavelengths. Also, we observe that RLMR outperforms QPS regardless of the number of wavelengths.

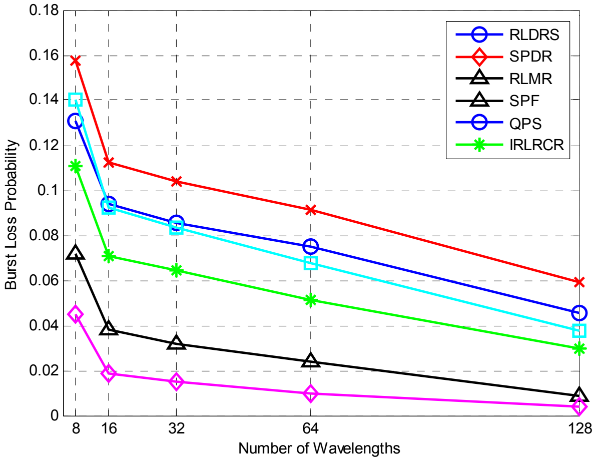


Figure 19. Loss probability vs. number of wavelengths on NSFNET.

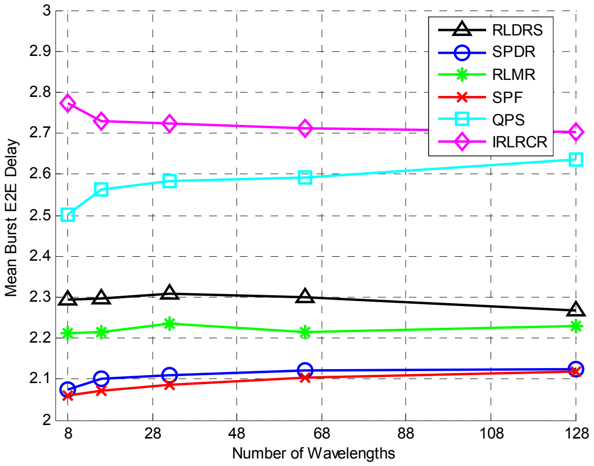


Figure 20. Mean burst end-to-end delay (number of hops) vs. number of wavelengths on NSFNET.

3.8. Conclusion and future work

In this paper, we proposed a novel reinforcement learning-based deflection routing scheme (RLDRS) which reduces effectively loss probability compared to Shortest Path Deflection Routing (SPDR). Also, we proposed an adaptive distributed reinforcement learning-based routing approach (RLMR) that reduces effectively burst loss probability compared to *Shortest Path First* routing (SPF) and to a recent and efficient multi-path routing algorithm called QPS. By combining RLMR and RLDRS, we defined an integrated reinforcement learning-based routing and contention resolution approach (IRLRCR). Simulation results show that RLMR, not only improves burst loss probability but also keeps burst end-to-end delay close to burst end-to-end delay of SPF, and considerably better than QPS. Besides, simulation results of RLDRS show its effectiveness in reducing loss probability compared to SPDR. In addition, the performance evaluation of IRLRCR shows that the combination of RLMR and RLDRS is successful since it reduces, effectively, loss probability compared to both RLMR and RLDRS, at the cost of a slight increase in mean burst end-to-end delay.

In the future, we plan to consider a multi-class traffic with IRLRCR in order to better exploit its tradeoff between burst loss probability and burst end-to-end delay. Moreover, we plan to develop congestion control mechanisms to deal with congestions that occur at very high loads.

Chapitre 4 :

Topology-aware wavelength partitioning for DWDM OBS networks: A novel approach for absolute QoS provisioning

Abdeltouab Belbekkouche, Abdelhakim Hafid, Mariam Tagmouti, Michel Gendreau

Abstract

Optical Burst Switching (OBS) is a promising switching technology for the next generation all-optical networks. An OBS network without wavelength converters and fiber delay lines can be implemented simply and cost-effectively using the existing technology. However, this kind of networks suffers from a relatively high burst loss probability at the OBS core nodes. To overcome this issue and consolidate OBS networks with QoS provisioning capabilities, we propose a wavelength partitioning approach, called Optimization-based Topology-aware Wavelength Partitioning approach (OTWP). OTWP formulates the wavelength partitioning problem, based on the topology of the network, as an Integer Linear Programming (ILP) model and uses a tabu search algorithm (TS) to resolve large instances efficiently. We use OTWP to develop an absolute QoS differentiation scheme, called Absolute Fair Quality of service Differentiation scheme (AFQD). AFQD is the first absolute QoS provisioning scheme that guarantees loss-free transmission for high priority traffic, inside the OBS network, regardless of its topology. Also, we use OTWP to develop a wavelength assignment scheme, called Best Effort Traffic Wavelength Assignment scheme (BETWA). BETWA aims to reduce loss probability for best effort traffic. To make AFQD adaptive to non-uniform traffic, we develop a wavelength borrowing protocol, called Wavelength Borrowing Protocol (WBP). Numerical results show the effectiveness of the proposed tabu search algorithm to resolve large instances of the partitioning problem. Also, simulation results, using ns-2, show that: (a) AFQD provides an excellent quality of service differentiation; (b) BETWA substantially decreases the loss probability of best

effort traffic to a remarkably low level for the OBS network under study; and (c) WBP makes AFQD adaptive to non-uniform traffic by reducing efficiently blocking probability for high priority traffic.

Keywords: Optical Burst Switching (OBS); Quality of Service (QoS); Admission Control; Linear Programming; Tabu Search; Wavelength Assignment; Fairness; Dense Wavelength Division Multiplexing (DWDM).

4.1. Introduction

Optical Burst Switching [18] is a promising switching technology for the next generation all-optical networks. It is considered as a tradeoff between Optical Circuit Switching (OCS) and Optical Packet Switching (OPS). OCS is easy to implement but suffers from poor bandwidth utilization and coarse granularity. OPS has a good bandwidth utilization and fine granularity but suffers from complex implementation because of the immaturity of the current technologies, such as optical buffers and ultra fast optical switches [18]. Hence, OBS is a good switching technology to benefit from the potential bandwidth that exists in optical fibers when used with Dense Wavelength Division Multiplexing (DWDM) technology. Indeed, theoretical research on OBS networks has reached the stage of prototypes in research laboratories [95, 96] and even commercial products (e.g., EtherBurst optical switch [97]). Hence, OBS networks could play an important role in metropolitan, access and local area optical networks.

In OBS networks, data packets with the same destination are aggregated in bursts of variable lengths at the ingress node, this is called *Burst Assembly*. After burst assembly, a *Control Packet* (also, called *Burst Header Packet*) is sent, using a dedicated control wavelength, from source to destination in order to reserve the required resources along a lightpath. This control packet is subject to *Optical/Electronic/Optical* (O/E/O) transformations at each core node (OBS switch) where it receives an appropriate processing. After a delay, called *Offset Time* (OT), the corresponding data burst is sent, on one of the data wavelengths, through the same lightpath (all-optically) without any buffering requirement inside the OBS network.

Wavelength contention is the main cause of burst losses in OBS networks. A wavelength contention occurs when two or more bursts intend to take the same output fiber, on the same wavelength, at the same time. Hence, a Quality of Service (QoS) scheme for OBS networks has to consider how to deal with wavelength contentions for each class of traffic. In fact, reducing loss probability for the OBS network is performed by reducing the level of wavelength contentions in the network. Moreover, in the context of multi-class traffic, reducing the loss probability of a given class of traffic can be performed by privileging this class in the case of contentions. It is worth noting that the *store-and-forward* based QoS schemes developed for electronic networks cannot be applied for OBS networks because of the lack of Random Access Memory (RAM) for optical networks.

In this paper, we consider absolute QoS provisioning for OBS networks. We propose a novel scheme, called *Absolute Fair Quality of service Differentiation scheme* (AFQD), which guarantees, for the first time, loss-free transmission for high priority bursts whatever the kind of the OBS network topology. AFQD is based on a wavelength partitioning approach, called *Optimization-based Topology-aware Wavelength Partitioning scheme* (OTWP), which uses Integer Linear Programming (ILP) to partition data wavelengths among the nodes in the network. Also, AFQD considers fairness among the users (nodes) of the network by allocating the same amount of bandwidth for high priority traffic to each user. Moreover, we propose a wavelength assignment scheme, called *Best Effort Traffic Wavelength Assignment scheme* (BETWA), which uses the wavelength partitioning approach OTWP to improve the performance of *Best Effort* traffic in terms of loss probability. To make AFQD adaptive to the case when high priority traffic pattern is non-uniform, we propose a Wavelength Borrowing Protocol (WBP) which aims to exploit non-utilized bandwidth while keeping the capabilities offered by AFQD (e.g., absolute QoS differentiation and fairness).

We consider an OBS network without wavelength converters and without Fiber Delay Lines (FDLs). This assumption is relevant since: (a) currently, wavelength conversion devices are complex, expensive, and not technologically mature; and (b) FDLs suffer from the lack of flexibility. Thus, the network under study can be implemented simply and cost-effectively using existing optical networks technology. Furthermore, this assumption allows measuring the performance improvement brought exclusively by our proposed schemes. We adopt Just Enough Time (JET) [18] protocol for resource reservation.

The remainder of this paper is organized as follows. In Section 4.2 we present related work. Section 4.3 presents a description of the proposed wavelength partitioning approach (OTWP), the exact ILP formulation of the wavelength partitioning problem and the proposed tabu search algorithm. Section 4.4, presents the proposed absolute QoS differentiation scheme (AFQD) and the proposed wavelength assignment scheme for best effort traffic (BETWA). Section 4.5 presents the proposed wavelength borrowing protocol (WBP). In Section 4.6, we present (1) numerical results that show the performance of the proposed tabu search algorithm and the ILP model resolution and (2) simulation results that show the performance of AFQD, BETWA and WBP. Finally, Section 4.7 concludes the paper.

4.2. Related work

4.2.1. Absolute QoS provisioning

In the literature, we find two kinds of QoS differentiation schemes for OBS networks: relative QoS differentiation [63, 70] and absolute QoS differentiation [74][75, 76, 79, 80]. Whereas absolute QoS guarantees, quantitatively, hard QoS requirements for high priority bursts, relative QoS just guarantees that high priority bursts will be served with higher quality (e.g., smaller loss probability) compared to low priority bursts. Notice that the main QoS parameter inside the OBS network is loss probability since data bursts are switched in the optical domain at each OBS switch without any queuing delay, especially, when Fiber Delay Lines (FDLs) are not used.

In [74-76, 79, 80], absolute QoS differentiation schemes are proposed. In [75], the proposed QoS differentiation scheme is based on a dynamic wavelength assignment approach where wavelengths are shared (dynamically) among different classes of traffic. The authors in [74] propose two schemes for absolute QoS provisioning: Early Dropping (ED) and Wavelength Grouping (WG). Early Dropping drops intentionally and probabilistically low priority bursts in order to maintain the level of loss probability of guaranteed bursts. Early dropping increases the loss probability for low priority traffic and the overall burst loss probability compared to the classless case. Wavelength Grouping (WG) provisions wavelengths for the guaranteed traffic and schedules bursts based on this provisioning mechanism. The authors in [76] propose a mechanism that guarantees a maximum loss probability for each guaranteed class of traffic. To do so, each core OBS node has to maintain traffic statistics for each class of traffic; a guaranteed burst can preempt a non-guaranteed burst (i.e., cancel its reservation) proactively according to a given preemption probability. The authors in [79] propose Reserve-and-Preempt Scheme (RPS) to improve the bandwidth provisioning of best effort traffic in the context of absolute QoS differentiation. In RPS, a metric, called Distance To Threshold (DTT), measures the difference between each priority guaranteed class loss rate and its pre-set loss rate threshold. Based on the value of DTT and using one of several proposed scenarios, a best effort burst can even preempt a priority guaranteed burst if the DTT of its class is: (a) big enough; and (b) the maximum among other priority guaranteed classes DTTs. The same authors in [80] propose several schemes which integrate RPS and wavelength grouping schemes in order to ensure that the highest priority guaranteed classes can be provided with their respective guaranteed loss rates. We can see that all of the above absolute QoS schemes consider a target value or a threshold for loss probability for high priority traffic.

4.2.2. Wavelength assignment

Despite its key role for data bursts transmission in OBS networks, wavelength assignment has not received much attention in the literature of OBS networks. This is due to the fact that most existing contributions assume that wavelength conversion capability is always present in OBS networks. However, this assumption is simplistic from both technical and economical points of view. Furthermore, this assumption could over-estimate the performance of OBS networks. Indeed, wavelength converters eliminate the wavelength continuity constraint (which stipulates that each data burst should be transmitted on the same wavelength from source to destination); a data burst could change its wavelength at any intermediate OBS node equipped with a wavelength converter. In addition, most of the authors use classical resource reservation policies in wavelength routed networks (e.g., First-Fit, Least Used and Random policies) for wavelength assignment in OBS networks. However, these policies are not adapted to OBS networks [61].

Almost all of the proposed wavelength assignment schemes for OBS networks can be characterized as feedback-based (adaptive) schemes [36, 61, 62]. The authors in [62] propose Priority-based Wavelength Assignment (PWA) algorithm. In PWA, each node in the OBS network maintains locally a priority value for each (wavelength, destination node) pair. Priorities are updated using feedbacks received after the transmission of data bursts; if the transmission was successful, priorities would be increased; otherwise, priorities would be decreased. To transmit a data burst, the node searches for a free wavelength in decreasing order of priorities. It is reported in [62] that PWA performs better than Random wavelength assignment policy (in terms of loss probability) under low traffic load; however, it performs only marginally better than Random policy under high traffic load. The authors in [61] propose two variants of PWA, namely, PWA-link and PWA-lambda. In PWA-link, each node associates a priority value to every (wavelength, link) pair (rather than (wavelength, node) pair in PWA) which results in finer granularity at the cost of more information exchange in the network. In PWA-lambda, each node associates a priority value to every wavelength; this makes PWA-lambda simpler and easier to implement but with worse performance compared to both PWA and PWA-link. Indeed, numerical results in [61] show that PWA-link is the best and PWA-lambda is the worst when comparing PWA, PWA-link and PWA-lambda. More recently, the authors in [36] propose another feedback-based wavelength assignment scheme, called Q-learning algorithm for Wavelength Selection (QWS). QWS is different from PWA in that it uses reinforcement algorithm (a Q-learning algorithm) to update priorities. The authors in [61] also propose a non-adaptive wavelength assignment scheme called First-Fit-TE where TE stands for traffic engineering. In First-Fit-TE, nodes are assigned start wavelengths depending on the traffic pattern and the fixed routing paths between nodes in the OBS network. Using First-Fit-TE, each node searches a free wavelength to transmit its data bursts starting from its start wavelength. The aim of First-Fit-TE is to improve the way First-Fit policy assigns a free wavelength. However, First-Fit-TE operation depends strongly on a priori known traffic pattern and (fixed) routing paths; this could limit its use (a) for OBS networks with dynamic (or even unknown) traffic patterns; and (b) when alternative routing is adopted. In this paper, we propose, for the first time, a wavelength assignment scheme (BETWA) which will be used to improve the performance of best effort traffic. BETWA uses the wavelength partitioning approach (OTWP) that formulates the partitioning problem as an Integer Linear Programming model (ILP); the objective is to assign a wavelength interval to each node based on the topology of the OBS network which rarely changes (in opposition to traffic patterns and routing paths that change more often).

4.3. Optimization-based topology-aware wavelength partitioning approach

In this Section, we present our wavelength partitioning scheme, called Optimization-based Topology-aware Wavelength Partitioning scheme (OTWP). First, we present a description of OTWP. Then, we present an exact formulation of wavelength partitioning in OTWP as an Integer Linear Programming (ILP) model. Finally, we present a tabu search algorithm to resolve efficiently the proposed ILP model.

4.3.1. OTWP description

The idea of OTWP is to allocate a number of wavelengths (one or more) to each OBS edge node in the network by considering the network topology. To this end, we model the OBS network as a graph $G(V, E)$ where V is the set of nodes ($|V| = N$) and E is the set of links ($|E| = M$). We suppose that each DWDM fiber link operates with W wavelengths where W is supposed to be bigger than N. If the number of wavelengths in each link is different, we can simply consider the number of wavelengths in the link with the minimum number of wavelengths or we can use a more sophisticated approach where each node is assigned a number of local wavelengths proportional to the number of wavelengths in its outgoing links; however, this is out of the scope of this paper. Also, additional optical fibers could be added to a link to increase its number of wavelengths (bandwidth capacity) if necessary; this ensures the scalability of OTWP. Hence, OTWP allocates a *wavelength interval* (i.e., a number of wavelengths) of size W / N to each node in the OBS network based on topological constraints, i.e., the closer two nodes to each other (i.e., the smaller the distance between the nodes), the more distant their allocated wavelength intervals from each other (i.e., the larger the distance between the intervals). We note that the size of wavelength intervals, represented by the ratio W / N, may be further increased by capturing edge nodes in the OBS network for which routing paths do not overlap (i.e., do not use any common link); in this case, bursts sent from these nodes will never contend among themselves, and hence, these nodes could use the same wavelength interval. For example, if we have h (two or more) edge nodes for which routing paths do not overlap, wavelength interval size is given by $W /(N-(h-1))$. The problem of increasing wavelength interval size is out of the scope of this paper; hence, in the rest of this paper, we consider that wavelength interval size is fixed to W / N.

For a given node $i \in \{0,..., N-1\}$, its allocated wavelength interval is denoted by $i' \in \{0,..., N-1\}$. The first wavelength in the wavelength interval i' is computed as follows:

$$St(i') = \begin{cases} \lfloor i' \cdot (W/N) \rfloor & \text{if } [i' \cdot (W/N) - \lfloor i' \cdot (W/N) \rfloor < 0.5] \\ \lceil i' \cdot (W/N) \rceil & \text{otherwise} \end{cases} \tag{4.1}$$

Thus, even if the size of wavelength intervals W/N is not always an integer value, we use floor and ceil functions to determine the start wavelength of each wavelength interval.

The distance between two wavelength intervals i' and j', denoted by $D(i', j')$, is calculated as follows:

$$D(i', j') = |i' - j'| \tag{4.2}$$

The distance between two nodes i and j in the network, denoted $d(i, j)$, is equal to the number of hops of the shortest path between nodes i and j.

Figure 21 shows an example of the operation of OTWP scheme using a linear topology of 4 nodes and 3 fiber links with 12 wavelengths on each link.

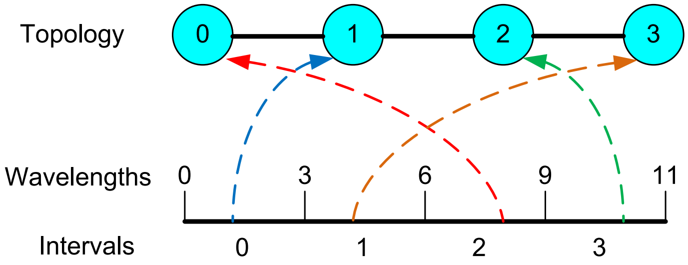


Figure 21. An example of wavelength partitioning using OTWP scheme.

We note that all of the nodes in this example are considered edge nodes (i.e., each node has its own traffic). In addition, node 1 and node 2 also play the role of core nodes (i.e., they can route traffic of the other nodes). In the optimal solution presented here (obtained by the resolution of the ILP model in subsection B using CPLEX 10.11 solver [98]), wavelength intervals 0 = [0, 2], 1= [3, 5], 2 = [6, 8] and 3 = [9, 11] are allocated to nodes 1, 3, 0 and 2, respectively. The distance between the wavelength interval of node 1 and the wavelength interval of its neighbor, node 2, is $D(0,3)=3$. Similarly, for nodes 0 and 1 we have $D(2,0)=2$ and for nodes 2 and 3 $D(3,1)=2$. These distances could be calculated for each pair of nodes in the network. We can expect that whenever these distances increase, wavelength contentions are less likely to occur, especially, when corresponding nodes are sending bursts on overlapping paths. For example, consider the case where node 1 and node 2 are sending bursts to node 3; if node 1 searches for an available wavelength first in its wavelength interval (i.e., wavelengths [0, 2]) and node 2 searches for an available wavelength in its wavelength interval (i.e., wavelengths [9, 11]), then wavelength contentions will occur less frequently than the case where both node 1 and node 2 search for available wavelengths using a classical approach, i.e., both nodes search for available wavelengths starting from wavelengths 0, 1, 2 and so on. Details on how the concept of distances between wavelength intervals is used practically are provided in Section 4.4.

4.3.2. Exact formulation

We formulate the above wavelength partitioning problem as a combinatorial optimization problem; in the following we present the model formulation.

Given:

$D[D_{i'j'}]$	Matrix of distances between wavelength intervals where $D_{i'j'}$ is calculated using Eq. (4.2).
$d[d_{ij}]$	Matrix of distances between nodes in the network where $d_{ij}=d_{ji}$ is the number of hops of the shortest path between node i and node j.
N	The number of nodes in the network and the number of wavelength intervals (each one of size W/N); we have one wavelength interval per node.

Variables:

$x_{ii'}$ A binary variable which takes the value 1 if wavelength interval i' is allocated to node i; 0 otherwise.

Objective:

$$Maximize \quad C = \sum_{i=0}^{N-1} \sum_{j=i+1}^{N-1} \sum_{i'=0}^{N-1} \sum_{j'=0}^{N-1} \frac{D_{i'j'}}{d_{ij}} x_{ii'} x_{jj'} \tag{4.3}$$

Subject to:

$$\sum_{i=0}^{N-1} x_{ii'} = 1 \quad i' = 0,\dots,N-1 \tag{4.4}$$

$$\sum_{i'=0}^{N-1} x_{ii'} = 1 \quad i = 0,\dots,N-1 \tag{4.5}$$

Bounds:

$x_{ii'} = 0,1 \quad i = 0,\dots,N-1 \quad i' = 0,\dots,N-1$

The objective function (4.3) maximizes the sum of ratios [distance between wavelength intervals/distance between nodes] over all the possible solutions of the partitioning problem. Indeed, we can consider C as a metric that measures the quality of any solution. The rational behind using the ratio $R_{ij} = \frac{D_{i'j'}}{d_{ij}}$ (for each pair of nodes i and j where $i < j$ and their allocated wavelength intervals i' and j', respectively) is that we have a maximization problem where we aim to increase the distance $D_{i'j'}$ whenever the distance d_{ij} is small. For example, let us consider again the topology shown in Figure 21, the metric $C = R_{01} + R_{02} + R_{03} + R_{12} + R_{13} + R_{23}$ is equal to $\frac{2}{1} + \frac{1}{2} + \frac{1}{3} + \frac{3}{1} + \frac{1}{2} + \frac{2}{1}$ in the optimal solution. We can see clearly that closer nodes have been allocated distant wavelength intervals and vice versa, which is our aim. Furthermore, any other solution that does not respect this property will not be optimal.

Constraints (4.4) state that each wavelength interval i' is allocated to exactly one node i in the network. Constraints (4.5) state that each node i in the network has exactly one allocated wavelength interval i'.

This model is similar to the Quadratic Assignment Problem (QAP) [99] with the difference that we maximize the metric C instead of minimizing a cost in a typical QAP. In fact, a standard formulation of QAP is given by:

$$\underset{\pi \in S_n}{Minimize} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{\pi(i)\pi(j)} b_{ij} \tag{4.6}$$

where S_n is the set of permutations of $\{1, 2, ..., n\}$ and each individual product $a_{\pi(i)\pi(j)} b_{ij}$ is the cost caused by assigning facility $\pi(i)$ to location i and facility $\pi(j)$ to location j.

Since the quadratic form in the objective function (4.3) makes the task of finding efficient resolution methods difficult, we formulate the problem as an Integer Linear Programming (ILP) model to benefit from its efficient resolution methods; thus, we define new variables ($y_{ij}^{i'j'}$) and constraints (4.7) with new bounds.

$y_{ij}^{i'j'}$	A binary variable which takes the value 1 if wavelength interval i' is allocated to node i and wavelength interval j' is allocated to node j; 0 otherwise.

$$2 y_{ij}^{i'j'} \leq x_{ii'} + x_{jj'} \tag{4.7}$$

with bounds:

$y_{ij}^{i'j'} = 0, 1 \; i = 0, ..., N-1 \; j = i+1, ..., N-1 \; i' = 0, ..., N-1 \; j' = 0, ..., N-1$

Constraints (4.7) forces variables $y_{ij}^{i'j'}$ to take the value 1 if both variables $x_{ii'}$ and $x_{jj'}$ take the value 1; this is always true because the coefficient of $y_{ij}^{i'j'}$, in the new objective function (4.8), is strictly positive in a maximization problem; otherwise, $y_{ij}^{i'j'}$ takes the value 0.

The new objective function becomes:

$$Maximize \quad C = \sum_{i=0}^{N-1} \sum_{j=i+1}^{N-1} \sum_{i'=0}^{N-1} \sum_{j'=0}^{N-1} \frac{D_{i'j'}}{d_{ij}} y_{ij}^{i'j'} \tag{4.8}$$

4.3.3. Complexity analysis

Even formulated as an ILP, our model remains a Quadratic Assignment Problem (QAP). It is known that QAPs are not only NP-hard but also remain among the hardest combinatorial optimization problems. In fact, the authors in [100] proved that QAP is NP-hard. This difficulty to resolve QAPs can be seen in their computational complexity. In our case, we can see that the set of feasible solutions is the set of all possible permutations of $\{0,...,N-1\}$ which is $n!$.

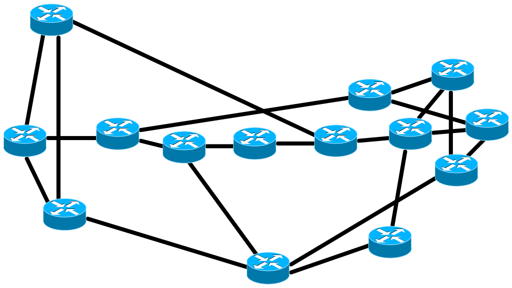


Figure 22. 14-nodes NSFNET topology.

Indeed, to have a good idea of the computational complexity of our ILP, we tried to resolve it using CPLEX 10.11 solver. Whereas the resolution time is reasonable for small instances, such as the network shown in Figure 21, it takes several days without returning the optimal solution for medium and large instances, such as 14-nodes NSFNET topology (Figure 22). Hence, a meta-heuristic approach that returns good solutions (rather than the optimal solution) in a reasonable time is clearly mandatory in this case.

4.3.4. Tabu search algorithm

The proposed tabu search algorithm aims to find, efficiently, a good solution to the wavelength partitioning problem in a reasonable time; it is inspired from the algorithm proposed in [101]. Tabu search has been proposed by Glover in 1986 [102]. This meta-heuristic searches the best feasible solution starting from the neighborhood of an initial solution. The process is repeated until a maximum number of iterations is reached. Tabu search avoids cycles by forbidding moves that take the current solution to a solution previously visited. In our case, these moves represent tabu pair exchanges between the nodes in the network; they are stored in a list called *tabu list*.

Let us define L as the function that returns the wavelength interval allocated to a given node; for instance, if wavelength interval i' is allocated to node i, then $L(i) = i'$. Also, let us denote a solution to the wavelength partitioning problem by a permutation P and the cost of this solution, defined in (4.8), by $C(P)$.

In this tabu search algorithm, a move represents an exchange of the assigned wavelength intervals between two nodes i and j, which means that if the current solution is a permutation P_1, then a new permutation P_2 is obtained by exchanging $L(i)$ and $L(j)$. The improvement when going from P_1 to P_2, denoted by $\Delta(P_1, P_2)$, is given by:

$$\Delta(P_1, P_2) = \sum_{k=0}^{N-1} \sum_{h=k+1}^{N-1} \frac{D_{L(k)L(h)}^{P_2}}{d_{kh}} - \sum_{k=0}^{N-1} \sum_{h=k+1}^{N-1} \frac{D_{L(k)L(h)}^{P_1}}{d_{kh}} \tag{4.9}$$

where D^{P_2} and D^{P_1} are the matrices of distances between wavelength intervals of solutions P_1 and P_2, respectively. The complexity of formula (4.9) is O(n^2). We reduce this complexity to O(n) by rewriting (4.9) as follows:

<u>*Lemma:*</u>

$$\Delta(P_1,P_2)=\sum_{\substack{k=0\\ k\neq i,j}}^{N-1}(D^{P_2}_{L(k)L(i)}-D^{P_1}_{L(k)L(j)})(\frac{1}{d_{ki}}-\frac{1}{d_{kj}}) \tag{4.10}$$

<u>*Proof:*</u>

Without loss of generality, we suppose that we have always $j>i$. For each node $k=0,...,N-1,\ k\neq i,j$, we distinguish three cases: (1) $k<i$; (2) $i<k<j$; and (3) $k>j$. For the three cases, we define $\Delta_1(P_1,P_2)$, $\Delta_2(P_1,P_2)$ and $\Delta_3(P_1,P_2)$, respectively, as follows:

$$\Delta(P_1,P_2)=\Delta_1(P_1,P_2)+\Delta_2(P_1,P_2)+\Delta_3(P_1,P_2) \tag{4.11}$$

where

$$\Delta_1(P_1,P_2)=\sum_{k=0}^{i-1}(\frac{D^{P_2}_{L(k)L(i)}}{d_{ki}}+\frac{D^{P_2}_{L(k)L(j)}}{d_{kj}}-\frac{D^{P_1}_{L(k)L(i)}}{d_{ki}}-\frac{D^{P_1}_{L(k)L(j)}}{d_{kj}}) \tag{4.12}$$

$$\Delta_2(P_1,P_2)=\sum_{k=i+1}^{j-1}(\frac{D^{P_2}_{L(i)L(k)}}{d_{ik}}+\frac{D^{P_2}_{L(k)L(j)}}{d_{kj}}-\frac{D^{P_1}_{L(i)L(k)}}{d_{ik}}-\frac{D^{P_1}_{L(k)L(j)}}{d_{kj}}) \tag{4.13}$$

$$\Delta_3(P_1,P_2)=\sum_{k=j+1}^{N-1}(\frac{D^{P_2}_{L(i)L(k)}}{d_{ik}}+\frac{D^{P_2}_{L(j)L(k)}}{d_{jk}}-\frac{D^{P_1}_{L(i)L(k)}}{d_{ik}}-\frac{D^{P_1}_{L(j)L(k)}}{d_{jk}}) \tag{4.14}$$

Since matrices D^{P_1}, D^{P_2} and d are symmetric, we determine using (4.11)-(4.14) that:

$$\Delta(P_1, P_2) = \sum_{\substack{k=0 \\ k\neq i,j}}^{N-1} \left(\frac{D^{P_2}_{L(k)L(i)}}{d_{ki}} + \frac{D^{P_2}_{L(k)L(j)}}{d_{kj}} - \frac{D^{P_1}_{L(k)L(i)}}{d_{ki}} - \frac{D^{P_1}_{L(k)L(j)}}{d_{kj}}\right)$$

$$= \sum_{\substack{k=0 \\ k\neq i,j}}^{N-1} \left(\frac{D^{P_2}_{L(k)L(i)} - D^{P_1}_{L(k)L(j)}}{d_{ki}} + \frac{D^{P_2}_{L(k)L(i)} - D^{P_1}_{L(k)L(j)}}{d_{kj}}\right)$$

$$= \sum_{\substack{k=0 \\ k\neq i,j}}^{N-1} (D^{P_2}_{L(k)L(i)} - D^{P_1}_{L(k)L(j)})\left(\frac{1}{d_{ki}} + \frac{1}{d_{kj}}\right) \qquad (4.15) \blacksquare$$

Algorithm 2. The Tabu Search (TS) algorithm.

```
Begin
  Step 0: Initialization
    Compute an initial feasible solution, denoted by $P_0$, using a construction method heuristic (see Algorithm 3 for more details);
    $best\_cost = C(P_0)$;
    int $k = 0$;
  Step 1: loop
    while ( $k < Max\_Iterations$)
      Find in the set of neighbours { $P_m$ obtained using a non tabu pair exchange on $P_k$ } of the current solution $P_k$, the best solution $P_{k+1}$ such that:
      $\Delta(P_k, P_{k+1}) = \underset{m}{Max}\, \Delta(P_k, P_m)$;
      add the pair exchanged to move from $P_k$ to $P_{k+1}$ to the tabu list;
      If ( $C(P_{k+1})$ > best-cost) then $best\_cost = C(P_{k+1})$;
      $k = k+1$;
    End while
End
```

Algorithm 2 shows the pseudo code of the proposed tabu search algorithm. The algorithm begins by the initialization step where an initial feasible solution has to be found.

This can be performed by generating a random feasible solution which can be a very bad solution. Hence, to improve the efficiency of the proposed tabu search algorithm, we propose a simple heuristic algorithm, which is a Construction Method (CM), to compute an initial feasible solution. Algorithm 3 shows the pseudo code of the proposed CM.

Algorithm 3. The Construction Method (CM) heuristic.

```
Initialization
    $S = \{\}$;
    $\overline{S} = \{V\}$;
    $Var\ i, i', j$;
Begin
    Find the node $i_0$ with the highest degree in the network;
    $S = \{i_0\}$;
    $\overline{S} = \overline{S} - \{i_0\}$;
    $L(i_0) = 0$;
    while ($\overline{S} \neq \emptyset$) do
    {
        $i = i$ the index of the node in $\overline{S}$ with the highest degree and which is a neighbor to one or more nodes in $S$;
        $L(i) = i'$ such that $\sum_{j \in S} \frac{D_{i'L(j)}}{d_{ij}}$ is maximal;
        $S = S \bigcup \{i\}$;
        $\overline{S} = \overline{S} - \{i\}$;
    }
End
```

CM uses a loop of N iterations (the number of nodes and wavelength intervals) where, at each iteration, a new node i is assigned a wavelength interval i' that maximizes the metric:

$$C' = \sum_{j \in S} \frac{D_{i'L(j)}}{d_{ij}} \tag{4.16}$$

where S is the set of nodes which have already been assigned a wavelength interval. After that, in step1, at each iteration k between 0 and *Max_Iterations*, the tabu search algorithm finds the best solution P_{k+1} in the neighborhood of the current solution P_k using a non-tabu pair exchange; if the cost of solution P_{k+1} is better than the best-cost found until now (denoted by *best_cost*), *best_cost* will take the value of the cost of P_{k+1}. The computational complexity of the loop in this algorithm is $O(n^2)$. The tabu list in our algorithm contains pairs (i, j) that cannot be exchanged in order to prevent cycles. We use a variable size tabu list, i.e., the size of the tabu list changes at each iteration of the algorithm. This approach improves the efficiency of the tabu search algorithm. Also, we implement this list as a two dimensional array which is updated in O(1).

4.4. Absolute fair QoS differentiation scheme

In this Section, we present the proposed Absolute Fair Quality of service Differentiation scheme (AFQD). First, we describe the operation of AFQD. Then, we present the proposed wavelength assignment scheme for best effort traffic (BETWA). Finally, we present a solution for the starvation problem of low priority traffic.

4.4.1. The operation of AFQD

AFQD uses the wavelength partitioning scheme OTWP, presented in Section 4.3, to provide absolute QoS differentiation. We assume that each wavelength interval assigned by OTWP to each edge node contains at least one wavelength. This assumption is very realistic since DWDM fiber links operate with tens to hundreds of wavelengths. Without loss of generality, we suppose that we have two classes of traffic: (a) Loss Sensitive (LS) traffic (e.g., mission critical applications traffic); and (b) Best Effort (BE) traffic. For simplicity, a burst belonging to LS traffic is called LS burst and a burst belonging to BE traffic is called BE burst. LS bursts have higher priority compared to BE bursts. We suppose that each OBS edge node has a burst assembly buffer for each destination OBS node and each class of traffic (i.e., LS or BE). Hence, an incoming packet from the client network (e.g., an IP network) is classified as LS or BE and forwarded to the appropriate assembly buffer according to: (1) its destination; and (2) its QoS requirements.

After the burst assembly phase, each LS burst is transmitted using one of the wavelengths of its source node's wavelength interval (we call these wavelengths: *local wavelengths*) along the OBS network with loss-free transmission guarantee. Thus, LS bursts from a given source node will never contend with LS bursts of the other nodes, and naturally, will never contend among themselves since they are transmitted from the same source of traffic. Consequently, there will be no wavelength contentions, and hence, no LS burst losses inside the network. Moreover, any classical wavelength assignment policy can be used to assign a local wavelength to a LS burst in the source node (e.g., First-Fit or Random). Differently from LS bursts, BE bursts can use any wavelength at the OBS source node; consequently, BE bursts can contend with LS bursts and BE bursts of the other nodes in the network.

Let us note that wavelength assignment at the OBS source node is decisive because of the absence of wavelength converters at the OBS core nodes. Thus, bursts (LS and BE) will use the same wavelength up to their OBS destination nodes. When two BE bursts are involved in a wavelength contention, one of the two bursts is dropped randomly; when a LS burst and a BE burst are involved in a wavelength contention, the LS burst is privileged to maintain the loss-free transmission guarantee; hence, even if a BE burst has already performed wavelength reservation, a LS burst can cancel this reservation and preempt the BE burst. If a LS burst cannot reserve a wavelength (in its OBS source node) even by preempting a BE burst, this will mean that the local wavelengths of this node are fully used by LS bursts; in this case, the LS burst can be stored (in the electronic domain) until a local wavelength becomes available, or it can be simply dropped. This can be seen as an admission control mechanism for LS traffic; a LS burst will never contend with another LS burst once it reserves a wavelength at its OBS source node. Figure 23 shows the operation of AFQD to schedule a LS burst at OBS source node. The test of LS traffic local wavelengths bandwidth utilization (and threshold U) is related to the starvation problem which is discussed in subsection 4.4.3.

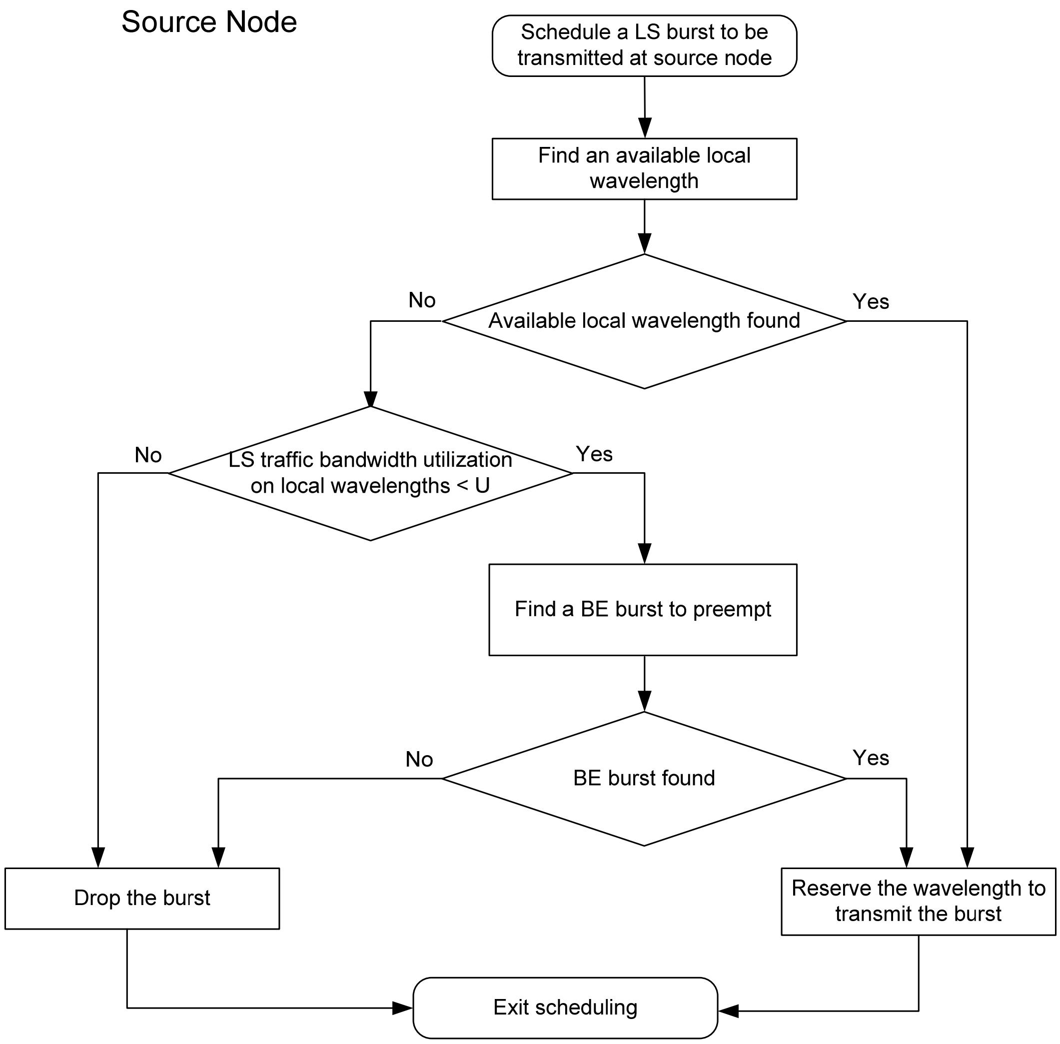


Figure 23. The operation of AFQD to schedule a LS burst at source node.

So far, we have explained how to guarantee loss-free transmission for LS bursts; the remaining challenge is to improve the performance of BE traffic as much as possible. For that, we propose a novel wavelength assignment scheme, called Best Effort Traffic Wavelength Assignment scheme (BETWA), to improve the performance of BE traffic.

4.4.2. Best effort traffic wavelength assignment scheme

BETWA is a general wavelength assignment scheme that can be used also in the classless case. However, we call it so to highlight its role in the scope of this paper (i.e., absolute QoS differentiation).

BETWA uses OTWP wavelength partitioning solution to find an available wavelength to transmit Best Effort (BE) bursts. Indeed, given a solution to the wavelength partitioning problem returned by OTWP, in a node i, BETWA searches an available wavelength to transmit a BE burst starting from the start wavelength of wavelength interval $(i'+1)$; the search process starts from wavelength interval $(i'+1)$ to alleviate contentions among LS bursts and BE bursts originating from the same node; also, since the closer the nodes to node i (e.g., one hop neighbors) the distant their wavelength intervals to i', wavelength contentions (and hence burst losses) among BE bursts will be considerably decreased. It is worth noting that a source node can assign any available wavelength (from the set of all wavelengths) to a BE burst; however, the order in which this available wavelength is searched is specific. Indeed, for a source node i with wavelength interval i', BETWA searches an available wavelength as follows: (1) starting from start wavelength of wavelength interval $(i'+1)$ to last wavelength of wavelength interval N; (2) from the last wavelength of wavelength interval $(i'-1)$ to the start wavelength of wavelength interval 0 (in the reverse direction); and finally (3) in the wavelength interval i' (from start wavelength to last wavelength). This order is adopted to benefit from the concept of distance between intervals. We conclude that OTWP maximizes (implicitly) the traffic isolation between the different nodes in the network. For example, in Figure 21, if BETWA is not used and a classical wavelength assignment is used instead (e.g., First-Fit), BE bursts originating from nodes 1 and 2 with destination 0 are more likely to contend on the link from node 1 to node 0. Figure 24 shows the order in which BETWA searches an available wavelength to schedule a BE burst at node i with wavelength interval i'. Figure 25 shows the operation of AFQD to schedule a BE burst at an OBS source node.

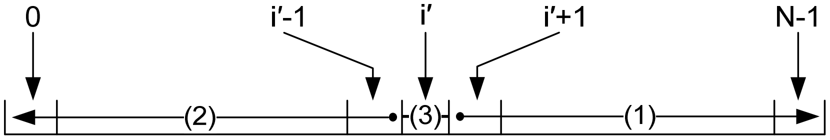


Figure 24. The order in which BETWA searches an available wavelength at node *i* with wavelength interval *i'*.

4.4.3. Starvation problem for BE traffic

Since LS bursts have the ability to preempt BE bursts, the BE bursts could suffer from the starvation problem when LS bursts use all of the available bandwidth in the network (i.e., each node uses fully its local wavelengths for LS traffic). Moreover, since the OBS network is considered, generally, as a transport network that serves client networks (e.g., IP networks, ATM networks, etc.), it is not always possible to predict the amount of LS traffic while planning the OBS network.

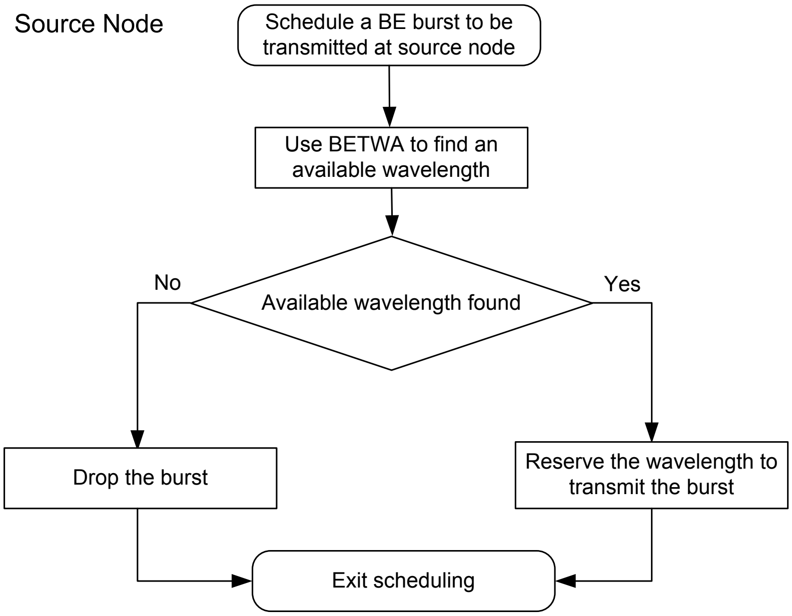


Figure 25. The operation of AFQD to schedule a BE burst at source node.

Hence, allowing unrestricted preemption of BE bursts makes it impossible to solve the problem of starvation. We propose to disable the preemption of BE bursts at an outgoing fiber link of a source node when the bandwidth utilization of LS traffic exceeds a predefined threshold, denoted by U, of the local wavelengths bandwidth of this node on this outgoing link (e.g., 80%). Thus, each OBS edge node has to monitor the bandwidth utilization of LS traffic at each one of its outgoing fiber links; whenever this bandwidth utilization exceeds the threshold U on a given outgoing fiber link, LS bursts cannot preempt BE bursts that have already reserved resources on this link. In this case, LS bursts are stored (in the electronic domain) until resources become available or simply dropped at the source node (admission control); this still preserves the loss-free transmission guarantee for LS traffic inside the OBS network. This scheme is simple since it operates only in the OBS edge nodes. Also, it allows LS traffic to use the whole capacity of the node's local wavelengths on a link when no (or negligible) BE traffic uses this link.

4.5. Wavelength borrowing protocol

So far, we have assumed that each OBS edge node has the same number of local wavelengths to send its LS traffic. However, an unbalanced (non-uniform) LS traffic in the network (i.e., some OBS edge nodes need to send more LS traffic than the capacity of their local wavelengths while other OBS edge nodes have no (or negligible) amount of LS traffic to send) may result in decreasing the performance of AFQD and thus degrading bandwidth utilization. To tackle this problem, we propose a wavelength borrowing scheme, called Wavelength Borrowing Protocol (WBP), which operates as follows: each OBS edge node keeps track of the amount of LS traffic sent on each one of its outgoing fiber links; whenever this amount exceeds the capacity of local wavelengths on an outgoing link, the node broadcasts a BORROW packet to the other edge nodes in the OBS network to request the use of a specific number of available unused wavelengths (depending on the amount of exceeding LS traffic) among the local wavelengths of other edge nodes.

Each edge node receiving a BORROW packet checks the utilization of its local wavelengths and answers this request only if it has available (non-utilized) wavelength (s); this is performed by reserving the available wavelengths (up to the requested number of wavelengths) and by sending a RESPONSE packet directly to the requesting node. If the node can provide the total number of requested wavelengths, it does not broadcast the received BORROW packet; however, if (a) it can only borrow a subset of the requested wavelengths; or (b) it cannot borrow any wavelength, it broadcasts the received BORROW packet to its neighbors. Each OBS edge node keeps track of the identifiers of the received BORROW packets to eliminate duplicates (i.e., when a BORROW packet is received twice or more). A requesting node can borrow from several nodes for the same request if none of the responding nodes can provide its requested number of wavelengths.

Upon receipt of RESPONSE packet, a requesting node responds the borrowing node by an ACCEPT packet if it accepts all or a subset of the proposed wavelengths; otherwise it sends a RELEASE packet to release all of the reserved wavelengths. The requesting node sends a RELEASE packet only if its demand in terms of wavelengths has been fulfilled by other nodes in the network. Upon receipt of ACCEPT packet, the borrowing node keeps the reservation of wavelengths included in the ACCEPT packet and makes available the wavelengths that were sent in RESPONSE packet and not included in ACCEPT packet. For better understanding, let us consider the following example: node1 sends BORROW packet requesting 3 wavelengths; node2, node3 and node4 send 3 RESPONSE packets including (wavelength1, wavelength2), (wavelength3, wavelength4), and (wavelength5, wavelength6), respectively. Node1 sends 2 ACCEPT packets including (wavelength1, wavelength2) and (wavelength3) to node2 and node3, respectively; it also sends RELEASE packet to node4. Node3 and node4 will make available wavelength4 and (wavelength5, wavelength6), respectively.

To ensure the stability of the wavelengths borrowing scheme and alleviate fluctuations, especially at the beginning of the network operation and when the traffic pattern is not yet in stable state (i.e., when the traffic pattern changes in very short time scale), WBP uses a timer to trigger any wavelength borrowing process; i.e., whenever LS traffic exceeds the capacity of local wavelengths on an outgoing link of a node, this node should wait a period of time T_{BORROW}, during which its LS traffic amount does not decrease under local wavelengths capacity, before sending a BORROW packet. Similarly, a borrowing node should wait a period of time T_{BORROW} during which its LS traffic amount does not increase above local wavelengths capacity, before sending a RESPONSE packet. Parameter T_{BORROW} can be set by the network operator.

When a borrowing node needs to recover its borrowed wavelengths (after an increase of its LS traffic), it sends a RECOVER packet to each node for which it has borrowed wavelengths. This node has to answer by a RELEASE packet to confirm that the borrowing node may again use its local wavelength (s).

Figure 26 shows the fields of WBP packets:

- *ID:* the identifier of the packet;
- *Packet type:* determines whether the packet is BORROW, RESPONSE, ACCEPT, RELEASE or RECOVER;
- *Source address:* the node that sends the packet;
- *Destination address:* the destination node of the packet; if this is a BORROW packet, the value of this field is a broadcasting address;
- *Number of wavelengths:* the number of wavelengths: (a) that are requested if the packet is BORROW; (b) that can be borrowed if the packet is RESPONSE; (c) that are actually borrowed if the packet is ACCEPT; and (d) that the borrowing node request to recover if the packet is RECOVER;
- *List of wavelengths:* used to determine the list of wavelengths that can be borrowed in a RESPONSE packet, the wavelengths actually borrowed in an ACCEPT packet and the list of wavelengths to recover in a RECOVER packet;

- *TTL (Time-To-Live):* used to prevent BORROW packets from being broadcasted indefinitely in the OBS network; the value of this field is decreased at each hop. An appropriate initial value of this filed should be equal to the diameter of the OBS network.

ID	Packet type	Source address	Destination address	Number of wavelengths	List of wavelengths	TTL

Figure 26. The fields of WBP packets

Figure 27 shows the interactions of WBP. Figure 27 (a) shows the interactions when a node requests wavelength (s) for its LS traffic and accepts all or a subset of the proposed wavelengths. Figure 27 (b) shows the interactions when a node requests wavelengths (s) and release all of the proposed wavelengths (if its request has been fulfilled by the other nodes). Figure 27 (c) shows the interactions when a borrowing node requests to recover its borrowed wavelengths (s).

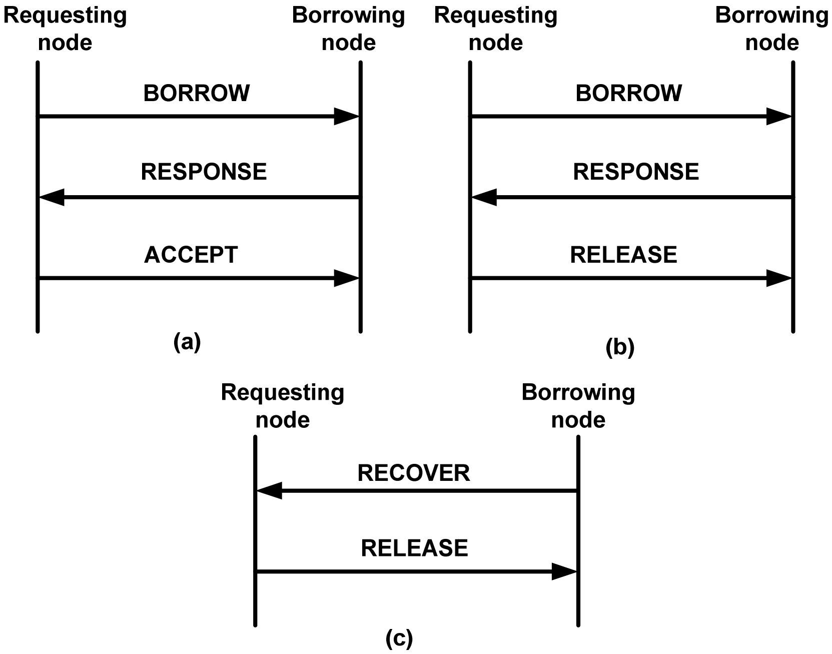


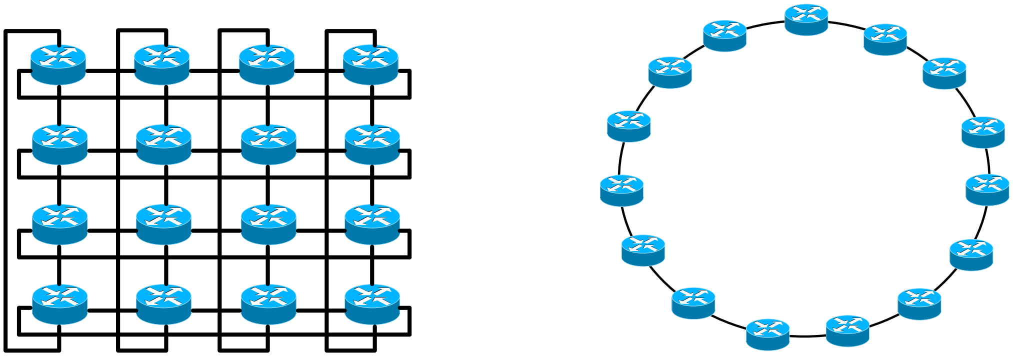
Figure 27. WBP interactions

It is worthwhile to note that the control packets of WBP, namely, BORROW, RESPONSE, ACCEPT, RELEASE and RECOVER will be sent on the control wavelength (s) of the OBS network. In addition, the fact that these packets have small sizes and the fact that they are sent sporadically (i.e., only when LS traffic on an outgoing link of a node exceeds its local wavelengths capacity) make them less likely to congest the control plan of the OBS network. Also, let us note that the loss of a WBP control packet (e.g., RELEASE packet) will need the use of timers and retransmission to make WBP reliable (i.e., TCP-like approach [103]); however; this is out of the scope of this paper since we suppose that the medium is reliable and that there is no congestion in the control plan.

4.6. Numerical results

In this Section, we present numerical results of CPLEX and Tabu Search algorithm and simulation results that show the performance of AFQD, WBP and BETWA. We use ns-2 simulator [93] and modules that implement OBS in ns-2 [94]. We consider three kinds of topologies: (1) mesh topologies represented by NSFNET with 14 nodes (Figure 22); (2) regular topologies represented by 4 x 4 nodes regular torus topology (Figure 28 (a)); and (3) ring topologies represented by a 15-nodes ring topology (Figure 28 (b)). Let us notice that we consider ring topologies because they have a primordial importance in metro and access optical networks.

We present NSFNET topology results in the majority of figures to alleviate presenting a huge number of figures, especially, when the general behavior is the same for all of the topologies. We assume that each single fiber link is bidirectional and all links have the same number of wavelengths. Each node in the network can generate, route and receive traffic (i.e., each node in the network is an edge and core node at the same time). Sources and destinations of traffic connections are generated randomly between any two nodes in the network, i.e., the traffic is dynamic and uniform; however, we use non-uniform traffic to measure the performance of WBP. The traffic load is expressed as the average traffic load per link and the capacity of a link is the sum of the capacities of all the wavelengths in this link. We use *Min Burst length Max Assembly Period* (MBMAP) algorithm for burst assembly [30]. We use exponential ON/OFF traffic and shortest path for routing. We consider loss probability which is the main performance metric in buffer-less OBS networks. Also, if a LS burst cannot be scheduled at an OBS source node, it is simply dropped (and not stored); the corresponding client networks are notified. Hence, LS bursts will not undergo additional delay at the border of the OBS network.



(a) (b)

Figure 28. (a) 4 x 4 regular topology and (b) 15-nodes ring topology.

All the following results have a confidence level of 95%.

4.6.1. Results of CPLEX and tabu search algorithm

To measure the quality of the proposed Tabu Search algorithm (TS) solutions, we performed experiments that compare the costs of its solutions to the costs of the solutions returned by CPLEX after a running time of 24 h (i.e., CPLEX has to return the solution, if any, at most after 24 h; this solution may be not optimal). Table 1 shows the average cost of solutions for random instances (topologies) of sizes from 10 to 100 nodes. We can see that CPLEX is unable to find solutions for instances of 50 nodes and more; this is because CPLEX needs an unrealistic amount of memory space to find even a first feasible solution for these instances. Moreover, these numerical results show that TS algorithm returns better solutions compared to CPLEX in few seconds to few minutes. Indeed, TS improves the solutions of CPLEX by 0.51% to 4.3% for instances between 10 nodes and 40 nodes. From computational time point of view, TS finds its solutions in less than 1 s for 10 nodes and 1576.5 s for 100 nodes.

4.6.2. Results of AFQD

The goal of these simulations is to measure the performance of AFQD. Since BETWA is a component of AFQD, we use it as a wavelength assignment scheme for BE traffic; the objective of using BETWA is to improve the performance of BE traffic when AFQD is offering absolute QoS provisioning to LS traffic.

Table 1. Comparison of Tabu Search (TS) and CPLEX results.

Number of nodes	CPLEX	TS
10	129.16	129.83
20	759.00	783.08
30	2191.92	2261.78
40	4955.67	5177.92
50	—	8620.80
60	—	15216.10
70	—	22185
80	—	30174
90	—	37827.20
100	—	38726.50

Simulation results of BETWA in the classless case (without QoS provisioning) are presented in subsection VI.C. We present: (a) *AFQD*: the overall loss probability for all of the bursts (i.e., blocking probability for LS bursts and loss probability for BE bursts); (b) *LS*: the blocking probability for LS traffic at the access of the OBS network (admission control blocking probability) since the loss probability of LS traffic is equal to zero inside the OBS network; and (c) *BE*: the loss probability for BE traffic. Also, to give an idea about the improvement brought by AFQD to the performance of OBS network, we plot the performance results of Latest Available Unscheduled Channel with Void Filling (LAUC-VF) [32] which is a good wavelength assignment algorithm that outperforms largely the classical wavelength assignment policies (e.g., First-Fit and Random). We note that since AFQD guarantees loss-free transmission to high priority traffic, it is not appropriate to compare it to existing QoS schemes for OBS networks which, in general, assume a maximum loss threshold for high priority traffic. We fix the value of the threshold U to 0.8, i.e., preemption is disabled when the bandwidth utilization of LS traffic reaches 80% of the bandwidth capacity of the local wavelengths of a node on an outgoing link. Unless stated otherwise, the amount of LS traffic that each node injects in the network is 1/N of the overall traffic that the node injects in the network (1/N represents the ratio [number of local wavelengths of each node (W/N)/total number of wavelengths in each fiber link (W)]). The remaining amount (i.e., [1-(1/N)]) is BE traffic.

Figure 29 shows the performance of AFQD compared to LAUC-VF. We can see that AFQD reduces effectively the loss probability of the OBS network. Indeed, whereas the mean loss probability of LAUC-VF (over all of the loads) is about 0.46, the mean loss probability of AFQD is about 0.29. In addition, whereas at load 50% loss probability of LAUC-VF is 0.47, it is 0.26 for AFQD. This proves that AFQD is not only able to provide absolute QoS differentiation for the OBS network, but also it can reduce its loss probability to a remarkable level for the network under study (i.e., without wavelength converters and FDLs). The same behavior is observed for regular and ring topologies (related figures are not presented here).

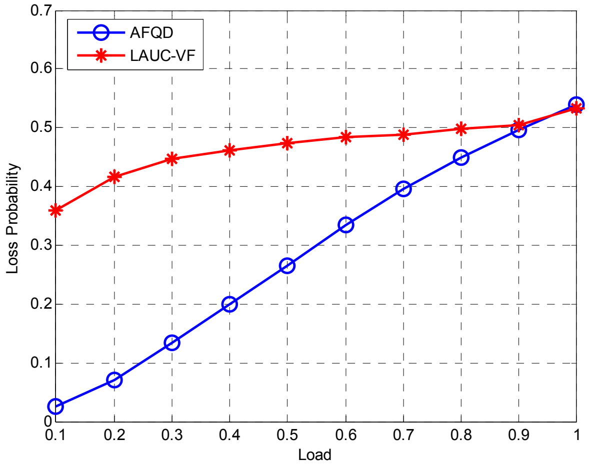


Figure 29. Loss probability vs. load for AFQD and LAUC-VF on NSFNET with 64 wavelengths.

Figure 30 shows the loss probability for BE traffic when varying the proportion of LS traffic (of the overall traffic in the network) from 0.5(1/N) to 2(1/N). We can see clearly that the proposed wavelength assignment scheme for BE traffic (BETWA) is efficient in reducing loss probability of BE traffic regardless of its proportion. Indeed, we can see that even at load 50%, the loss probability for BE traffic is in the order of 0.2. Also, we can see that BE loss probability slightly increases when the proportion of LS traffic increases which was expected.

Figure 31 shows the blocking probability of LS traffic at the source OBS edge nodes when varying the proportion of LS traffic. This blocking could happen when the LS traffic bandwidth utilization of the local wavelengths of a node on an outgoing link exceeds the threshold U (80% in these simulations). We can see that this blocking probability is at most in the order of 10^{-2} when the proportion of LS traffic is (1/N) (i.e., LS traffic uses 100% of local wavelengths bandwidth capacity); this is comparable to the high priority traffic loss target in other QoS schemes (e.g., 10^{-2} in [76] and $2x10^{-3}$ in [79]) where losses are allowed inside the OBS network. Also, we can see that this blocking probability increases when the proportion of LS traffic increases; for example, when the proportion of

LS traffic is 2(1/N) (i.e., twice the capacity of local wavelengths), LS bursts blocking probability may reach 10^{-1} at load 50%.

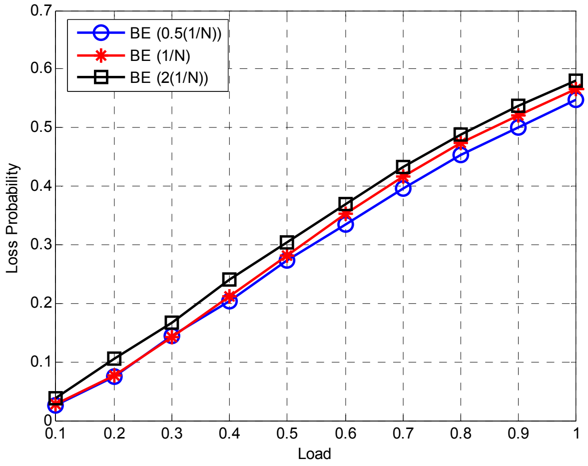


Figure 30. BE traffic loss probability vs. load on NSFNET with 64wavelengths.

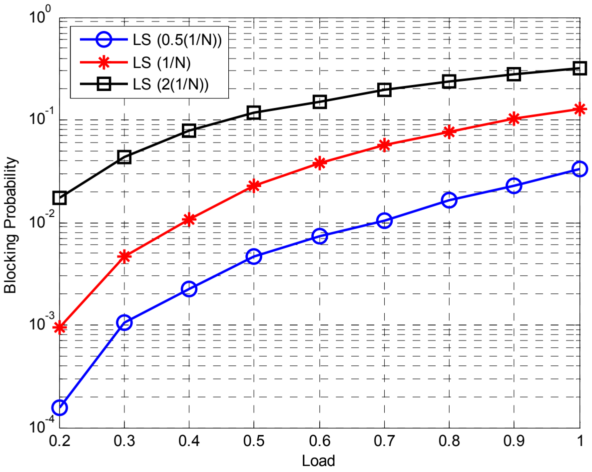


Figure 31. LS traffic blocking probability on NSFNET with 64wavelengths

From Figure 30 and Figure 31 we observe that blocking probability of LS traffic is significantly lower than the loss probability of BE traffic. This proves that AFQD successfully provides QoS differentiation among LS traffic and BE traffic.

Figure 32 shows loss probability of AFQD and LAUC-VF when varying the number of wavelengths from 32 to 160 using NSFNET topology and fixing the traffic load at 50% of the capacity of the network at each value of the number of wavelengths (obviously, the amount of traffic corresponding to 50% traffic load increases when the number of wavelengths increases). We can see that whatever the number of wavelengths, AFQD outperforms significantly LAUC-VF. Also, we observe that whereas loss probability of LAUC-VF increases when the number of wavelengths increases, loss probability of AFQD slightly decreases. Indeed; this behavior was expected since the size of each wavelength interval (the number of local wavelengths) in OTWP becomes larger when the number of wavelengths increases; this increases the capability of traffic isolation amongst the nodes in the OBS network. Also, this proves that AFQD is more efficient when using DWDM technology where each fiber link operates with a large number of wavelengths. Furthermore, the increase of loss probability of LAUC-VF when increasing the number of wavelengths is explained by the fact that LAUC-VF is unable to exploit the increase of the number of wavelengths; indeed, it performs wavelength assignment in each node using only local information and without considering the traffic of the other nodes in the network.

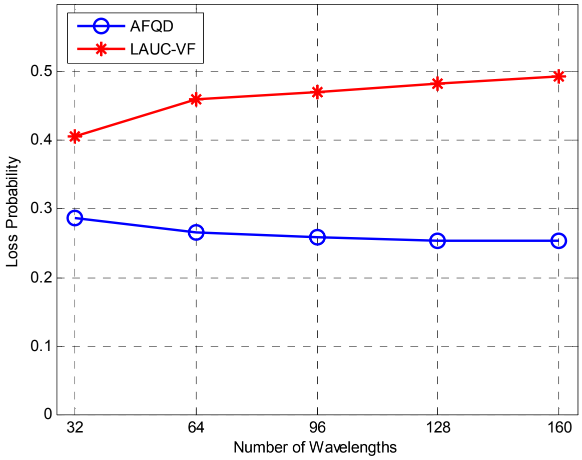


Figure 32. Loss probability of AFDQ and LAUC-VF on NSFNET when varying the number of wavelengths.

Figure 33 shows loss probability of BE and LAUC-VF and blocking probability of LS using 4 x 4 regular topology and 64 wavelengths. We can see that the trends are similar to the case of NSFNET topology.

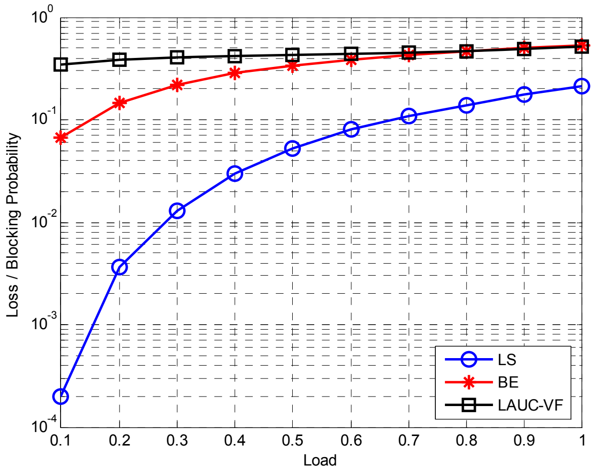


Figure 33. Blocking probability of LS and loss probability of BE and LAUC-VF on 4 x 4 regular topology with 64 wavelengths.

Figure 34 shows loss probability of BE and LAUC-VF and blocking probability of LS using 15-nodes ring topology. Here again, the results are similar to those of NSFNET and regular topologies. In addition, we observe that LS blocking probability performance is better compared to NSFNET and regular topologies. In fact, the blocking probability of LS traffic is as low as 10^{-4} at 20% load. We conclude that AFQD performs better in OBS ring networks which are, usually, used for optical Metropolitan Area Networks (MANs)

This proves that AFQD is able to provide QoS differentiation and reduce significantly loss probability regardless of the topology of the OBS network.

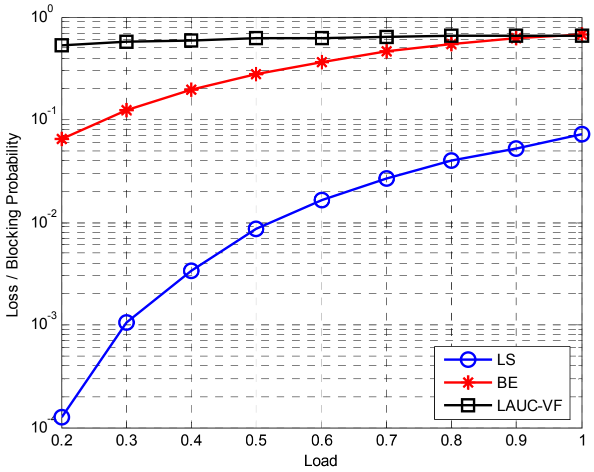


Figure 34. Blocking probability of LS and loss probability of BE and LAUC-VF on 15-nodes ring topology with 64 wavelengths.

4.6.3. Results of BETWA

As stated in section 4.4.2, BETWA can be seen as a separate contribution in the context of wavelength assignment for classless OBS networks. Thus, we perform simulations to measure the performance of BETWA in the general case (classless case). We compare BETWA to Latest Available Unscheduled Channel with Void Filling (LAUC-VF) [32] and Q-learning algorithm for Wavelength Selection (QWS) [36] (see Section II). LAUC algorithm uses the latest available wavelength, i.e., the available wavelength with the minimum time interval between the end time of the last scheduled burst and the start time of the burst to be scheduled. Void Filling (VF) has been introduced to allow scheduling bursts in voids (i.e., between two already scheduled bursts) and hence, reduces fragmentation and increases wavelength utilization. The motivations behind comparing BETWA to LAUC-VF and QWS are: (a) LAUC-VF is a good wavelength assignment algorithm that outperforms largely the classical wavelength assignment policies (e.g., First-Fit and Random); and (b) QWS is a recent scheme for wavelength assignment in OBS networks that outperforms PWA-link and First-Fit-TE [61] (see Section 4.2); hence, by comparing BETWA to QWS, we are comparing it indirectly to PWA-link and First-Fit-TE. For QWS parameters, we use the same values reported in [36].

Figure 35 shows loss probability with NSFNET topology and 64 wavelengths. We can see clearly that BETWA outperforms LAUC-VF and QWS. In fact, the mean burst loss probability (over all of the loads) of BETWA, QWS and LAUC-VF are 0.27, 0.53 and 0.46, respectively. We conclude that BETWA reduces effectively loss probability. Also, we observe that QWS is better than LAUC-VF only for low traffic loads (i.e., under 20%).

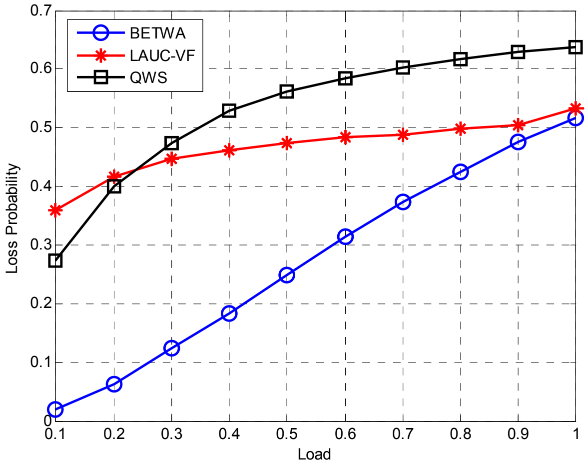


Figure 35. BETWA: loss probability vs. load with NSFNET and 64 wavelengths.

To measure the impact of varying the number of wavelengths in each fiber link on the performance of BETWA, we consider NSFNET topology and vary the number of wavelengths from 32 to 160 while fixing the traffic load at 50% of the network capacity when varying the number of wavelengths (Figure 36). We can see that, whatever the number of wavelengths, BETWA outperforms the other schemes. Also, we observe that whenever the number of wavelengths increases, the performance of BETWA becomes slightly better; this is because the size of wavelength intervals increases and because BETWA exploits efficiently the additional bandwidth offered by the increase of the number of wavelengths; this is not the case of (a) LAUC-VF which performs wavelength assignment based on local information in each node; and (b) QWS which performs wavelength assignment in each node by probing the state of the network using a reinforcement learning approach.

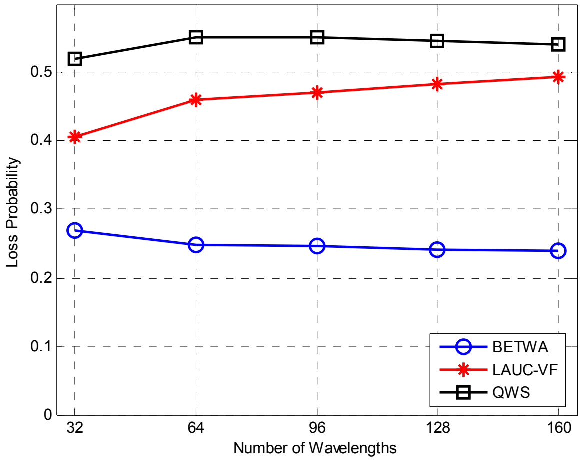


Figure 36. BETWA: loss probability vs. number of wavelengths with NSFNET topology when load is fixed to 50%.

With regular topology (Figure 37) and 64 wavelengths, we observe the same trends as with NSFNET topology (Figure 35). However, in ring topology (Figure 38), QWS is better than LAUC-VF whatever the value of traffic load in the network.

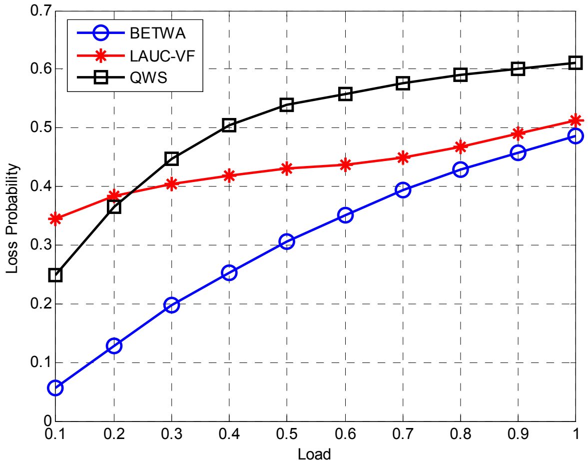


Figure 37. BETWA: loss probability vs. number of wavelengths with 4 x 4 nodes regular topology and 64 wavelengths.

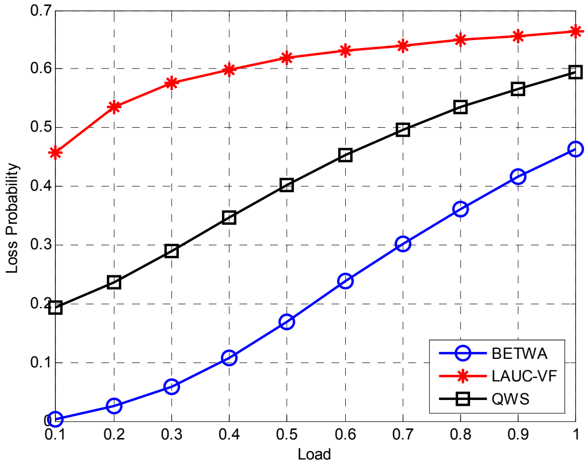


Figure 38. BETWA: loss probability vs. number of wavelengths with ring topology and 64 wavelengths.

4.6.4. Results of WBP

To show the capability of WBP to reduce blocking probability of LS traffic, we perform simulations using non-uniform LS traffic where only half of the network nodes send LS traffic. We set the proportion of LS traffic to twice the capacity of local wavelengths in each sending node (i.e., 2(1/N)).

Figure 39 shows blocking probability of LS traffic with WBP and without WBP on NSFNET topology with 64 wavelengths. We can see clearly that WBP is able to reduce effectively blocking probability of LS traffic, especially, when the traffic pattern is non-uniform and when some nodes are not fully using their local wavelengths. Indeed, the mean blocking probability (over all of the loads) of LS traffic when using WBP is 0.07 and the mean blocking probability without using WBP is 0.32.

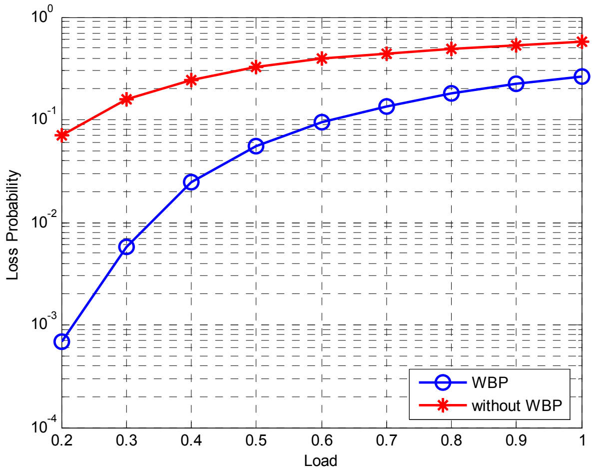


Figure 39. Blocking probability of LS with and without WBP on NSFNET and 64 wavelengths.

The same trends as NSFNET are observed for regular topology (Figure 40); when WBP is used, it reduces effectively blocking probability when the LS traffic pattern is non-uniform.

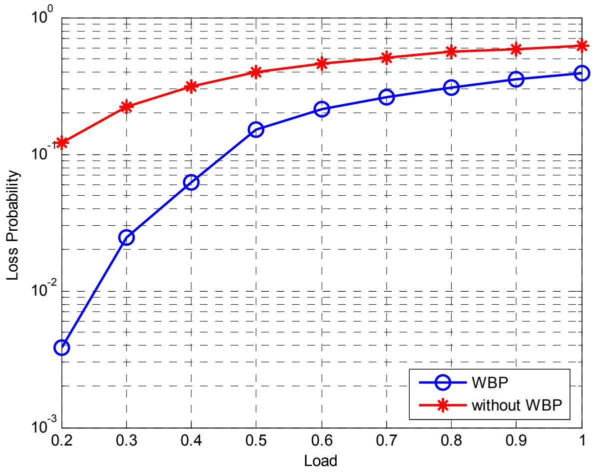


Figure 40. Blocking probability of LS with and without WBP on regular topology with 64 wavelengths.

Figure 41 shows the performance of WBP with ring topology. We observe that, in ring topology, WBP performs better compared to NSFNET and regular topologies. In fact, at 40% load, WBP is able to reduce LS traffic blocking probability by orders of magnitude, i.e., from 10^{-1} to 10^{-4}.

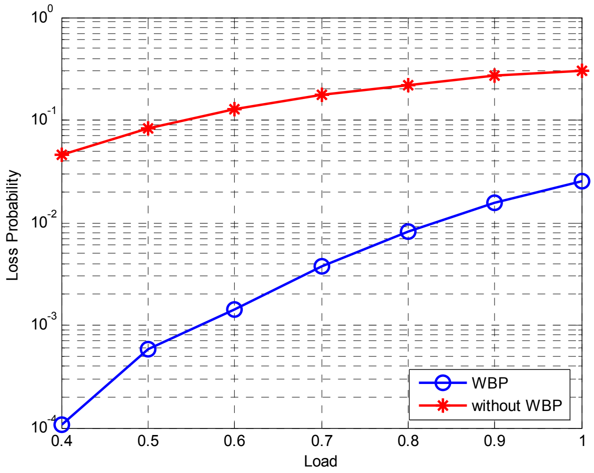


Figure 41. Blocking probability of LS with and without WBP on ring topology and 64 wavelengths.

4.7. Concluding remarks

In this paper, we have proposed an absolute QoS differentiation scheme for OBS networks (AFQD). AFQD is based on a wavelength partitioning scheme (OTWP) which models the wavelength partitioning problem as an Integer Linear Programming (ILP) model and uses a tabu search algorithm to resolve it efficiently. AFQD guarantees loss-free transmission inside the OBS network for high priority traffic (Loss Sensitive (LS) traffic) whatever the kind of the OBS network topology. In addition, AFQD uses a novel wavelength assignment scheme (BETWA), based on OTWP, to improve the performance of Best Effort (BE) traffic in terms of loss probability. Also, BETWA can be used as an efficient wavelength assignment scheme in the context of classless OBS networks. To make AFQD adaptive to non-uniform traffic patterns, we proposed a wavelength borrowing protocol (WBP). Simulation results did show that: (1) AFQD is effective to provide absolute QoS differentiation and to guarantee loss-free transmission for LS traffic; (2) BETWA decreases, significantly, loss probability for BE traffic (with AFQD) and overall traffic (without AFQD) to a remarkably low level for the OBS network under study; (3) WBP is capable of reducing, effectively, blocking probability of LS traffic with non-uniform traffic.

Chapitre 5 :

Path-based QoS Provisioning for Optical Burst Switching Networks

Abdeltouab Belbekkouche, Abdelhakim Hafid, Michel Gendreau and Mariam Tagmouti

Abstract

Optical Burst Switching (OBS) networks are candidates to play an important role in next generation optical networks where Quality of Service (QoS) provisioning is an essential feature. In this paper, we investigate the ability of OBS networks to guarantee loss-free transmission inside the network for guaranteed bursts. More specifically, we propose a QoS approach, called Path-based QoS Provisioning (PQP), to provide absolute QoS provisioning for OBS networks. PQP relies on: (a) routing and wavelength assignment to establish, whenever possible, non-overlapping paths between each pair of OBS edge nodes; and (b) a synchronization scheme to guarantee QoS with efficient sub-wavelength resource utilization when the solution in (a) contains overlapping paths because of the limited number of wavelengths. For (a) we propose a routing and wavelength assignment approach, called OBS Routing and Wavelength Assignment (OBSRWA), which uses an efficient Integer Linear Programming (ILP) model to determine routing paths and a tabu search algorithm to assign wavelengths to these paths. For (b) we propose a path synchronization scheme, called Path-based Synchronous Transmission scheme (PST). PST synchronizes the transmissions in each set of overlapping paths while maximizing the capacity of each path to transmit guaranteed traffic and guaranteeing fairness when allocating bandwidth to conflicting paths; this is performed using efficient ILP formulations. To improve the performance of best effort traffic and preserve statistical multiplexing and high resource utilization of the OBS network, we propose a wavelength selection scheme, called Path-based Best Effort Wavelength Selection scheme (PBEWS), to send best effort bursts.

Simulation results using ns-2 simulator show that PQP successfully provides absolute QoS provisioning for guaranteed traffic and improves significantly the performance of best effort traffic.

Keywords: Optical Burst Switching (OBS); QoS Provisioning; Routing; Wavelength Assignment; Linear Programming; Tabu Search, Synchronization.

5.1. Introduction

Optical Burst Switching is a switching technology that was proposed as a candidate transport solution for next generation optical Internet [18]. OBS is considered as a tradeoff between Optical Circuit Switching (OCS) and Optical Packet Switching (OPS). For example, OBS networks could be a better choice than OCS networks for metro-core networks [6] and they may play the role of edge optical networks to reduce the electronic-grooming requirements at the edge-core interface [104]. Hence, the OBS paradigm is candidate to play an important role in Next Generation Networks (NGN) framework for which QoS provisioning is an essential feature [105].

The lack of efficient optical buffers makes the task of designing Quality of Service (QoS) provisioning mechanisms for OBS networks less straightforward compared to the case of electronic networks for which QoS mechanisms are based on the *store-and-forward* concept. Indeed, the enhanced buffering capabilities of electronic networks (e.g., IP networks) allows QoS provisioning using per-class queuing, buffer management and advanced scheduling policies which are not possible to apply at the core of the OBS network because of the lack of efficient buffers. In addition, any QoS provisioning mechanism for OBS networks has to consider how to deal with wavelength contentions (i.e., when two or more bursts intend to take the same output fiber, on the same wavelength, at the same time) for each class of traffic. In fact, reducing loss probability for the OBS network is performed by reducing the level of wavelength contentions in the network.

Moreover, in the context of multi-class traffic, reducing the loss probability of a given class of traffic can be performed by privileging this class when contentions occur.

In this paper, we investigate the ability of OBS networks to provide absolute QoS provisioning in terms of loss probability at end-to-end path level (i.e., path-based QoS provisioning). More specifically, we investigate the ability of OBS networks to provide loss-free transmission inside the OBS networks for guaranteed traffic. This study is motivated by the fact that establishing non-overlapping paths between each pair of OBS edge nodes requires, generally, a number of wavelengths which is far smaller than the widely believed $O(N^2)$ where N is the number of edge nodes; this is because the topology of the OBS network is less likely to be a complete graph. In fact, if the problem of assigning wavelengths to paths is modeled as a vertex coloring problem, the number of required colors (wavelengths) is the chromatic number $\chi(G')$ of the conflict graph of the OBS network G' given a routing paths configuration where each vertex in the conflict graph represents a path and an edge exists between two vertices if the corresponding paths overlap (share at least a link on the same wavelength). Table 2 shows the required number of wavelengths ($\chi(G')$) to establish completely non-overlapping paths for twelve of the largest SNDlib instances [106] when shortest path routing is used. We can see clearly that the number of required wavelengths is far smaller than $N(N-1)$. Furthermore, this number can be further decreased by finding a routing paths configuration which corresponds to the conflict graph with the minimum chromatic number (see Section 5.6.1); this reduces the number of required wavelengths to color the conflict graph, and hence, makes our approach to provide path-based QoS provisioning a viable solution for most OBS networks' topologies. In the case where some paths still overlap because of the limited number of wavelengths in each fiber link, we propose to use *limited scale synchronization* to guarantee loss-free transmission in the overlapping paths while maximizing the amount of guaranteed traffic that can be carried by each path, increasing resource utilization and guaranteeing fairness when allocating bandwidth to different paths.

Table 2. The number of required wavelengths to establish non-overlapping paths using shortest path routing for twelve of SNDlib instances.

Instance	**Number of nodes (*N*)**	**Number of links**	***N(N-1)***	**Number of required wavelengths**
nsfnet (nobel-us)	14	21	182	15
atlanta	15	22	210	26
newyork	16	49	240	13
France	25	45	600	51
janos-us	26	84	650	88
cost266	37	57	1332	162
giul39	39	172	1482	77
janos-us-ca	39	122	1482	184
pioro40	40	89	1560	144
germany50	50	88	2450	236
zib54	54	81	2862	336
ta2	65	108	4160	433

The contributions in this paper are as follows:

- A novel approach for routing and wavelength assignment problem including: **(a)** an efficient ILP formulation for the routing paths problem; **(b)** an ILP formulation for the wavelength assignment problem; and **(c)** a tabu search algorithm to resolve the wavelength assignment problem for large instances;

- A path synchronization scheme for overlapping paths which uses: **(d)** a synchronization protocol that eliminates the need for Fiber Delay Lines (FDLs) to synchronize the transmissions of different paths; and **(e)** efficient ILP models to maximize the amount of guaranteed traffic that can be carried by each path; and

- **(f)** A wavelength selection scheme which improves the performance of best effort traffic and keeps the characteristics of statistical multiplexing gain and high resource utilization of the OBS network.

We consider an OBS network without wavelength converters and without FDLs. This assumption is relevant since: (a) currently, wavelength conversion devices are complex, expensive, and not technologically mature; and (b) FDLs suffer from the lack of flexibility. Thus, the network under study can be implemented simply and cost-effectively using the existing optical networks technology. Furthermore, this assumption allows measuring the performance improvement brought exclusively by our proposed approach. Besides, we adopt Just Enough Time (JET) [18] protocol for resource reservation.

The remainder of this paper is as follows. Section 5.2 presents related work on QoS provisioning and synchronization for OBS networks. Section 5.3 presents the proposed QoS provisioning approach (PQP) and the wavelength selection scheme (PBEWS). Section 5.4 introduces the proposed routing and wavelength assignment approach (OBSRWA). Section 5.5 presents the proposed synchronization approach (PST). Section 5.6 presents numerical results and Section 5.7 concludes the paper.

5.2. Related work

5.2.1. QoS provisioning

In the literature, we find two kinds of QoS differentiation schemes for OBS networks: relative QoS differentiation [63, 70] and absolute QoS differentiation [74][75, 76, 79, 80]. Whereas absolute QoS guarantees, quantitatively, hard QoS requirements for high priority bursts, relative QoS just guarantees that high priority bursts will be served with higher quality (e.g., smaller loss probability) compared to low priority bursts. Notice that the main QoS parameter inside the OBS network is loss probability since data bursts are switched in the optical domain at each OBS switch without any queuing delay, especially, when Fiber Delay Lines (FDLs) are not used.

In [74][75, 76, 79, 80], different absolute QoS differentiation schemes are proposed based on dynamic wavelength provisioning to different classes of traffic [75], early dropping and wavelength grouping [74], recording traffic statistics for each class of traffic and using preemption to guarantee a maximum loss probability threshold for guaranteed bursts [76] and improving the bandwidth provisioning of best effort traffic by using Distance-To-Threshold (DTT) metric [79, 80]. All of these schemes propose a maximum loss probability threshold for high priority traffic, and hence, allow high priority bursts losses inside the OBS network. In this paper, we are interested in providing loss-free transmission inside the OBS network for high priority traffic. More recently, we have proposed [107] a wavelength partitioning approach that allocates a number of wavelengths (one or more) to each OBS edge node to send its guaranteed traffic with loss-free transmission guarantee inside the OBS network; we have assumed that fiber links are operating with DWDM multiplexing technology and that the number of wavelengths is bigger than the number of nodes in the network. In this paper, we investigate the capability of the OBS network to provide loss-free transmission inside the OBS network at path level without limitations/assumptions on the number of wavelengths in each fiber link and the topology of the OBS network.

5.2.2. Synchronization in OBS networks

Synchronization in OBS can be thought as Time Division Multiplexing (TDM) over Wavelength Division Multiplexing (WDM) where bursts of predefined fixed duration, called timeslots, are switched in the time domain rather than the wavelength domain. This concept has attracted an increasing interest during the last several years because of its ability to improve the performance and resource utilization of the network compared to asynchronous (classical) OBS [10-16]. Ramamirtham *et al.* [10] were the first to introduce the concept of time sliced (or time slotted) OBS. They proposed to (1) send bursts in fixed duration timeslots and to synchronize all of the incoming timeslots at the input of an OBS switch using synchronizers (fiber delay lines); and (2) perform the switching of timeslots in the time domain using devices called Optical TimeSlot Interchangers (OTSI); this architecture eliminates the need for wavelength converters in the OBS network. Sheeshia *et al.* [11] proposed a protocol, called Synchronous Optical Burst Switching (SOBS), to support synchronous services, such as SONET/SDH; they proposed several solutions regarding path signaling, periodic reservations and burst framing. Particularly, a two way reservation scheme is proposed for synchronous traffic. Ozturk *et al.* [13] studied the performance of slotted OBS using discrete-time Markov chain (DTMC) based framework to determine the loss probabilities in the cases with and without QoS differentiation. For the case of QoS differentiation, they have considered two mechanisms: priority scheduling-based QoS differentiation and Offset Time-based QoS differentiation. More recently, Jeong *et al.* [14] proposed to use timeslot allocation to offer loss-free transmission inside the OBS network. To this end, a Tree-based Slot Allocation (TSA) algorithm is proposed where in each *superframe*, destined to a destination node input link, timeslots in different positions are allocated to nodes to alleviate contentions; also, they proposed a multiplexing optimization of superframes technique to reduce the required number of wavelengths.

In all these existing schemes [10-16], the whole OBS network is synchronized using fiber delay lines at the input of switches. In this paper, we use synchronization only when necessary and at a limited scale, i.e., only to synchronize the transmissions of overlapping paths. Also, we propose a synchronization protocol which allows synchronization without the need for fiber delay lines.

5.3. Path-based QoS provisioning

In this section, we describe the details of PQP. First, we present PQP where we distinguish two cases: (a) the number of available wavelengths, the number of paths and the routing paths configuration allow to establish non-overlapping paths between each pair of edge nodes; we call this case asynchronous PQP; and (b) the limited number of wavelengths makes that some paths overlap (i.e. use the same wavelength on at least one link); we call this case synchronous PQP; we explain how PQP provides QoS provisioning in each case. Then, we present a wavelength selection scheme, called Path-based Best Effort Wavelength Selection (PBEWS), to send best effort bursts. PBEWS improves the performance of best effort traffic and preserves the characteristics of statistical multiplexing and high resource utilization of the OBS network.

5.3.1. Overview

PQP aims to provide absolute QoS provisioning to high priority traffic classes. Without loss of generality, we suppose that we have two classes of traffic: (a) Loss Sensitive (LS) traffic (e.g., mission critical applications traffic); and (b) Best-Effort (BE) traffic. For simplicity, a burst belonging to LS traffic is called LS burst and a burst belonging to BE traffic is called BE burst. LS bursts have higher priority compared to BE bursts. Based on the number of available wavelengths in each fiber link, the number of paths in the OBS network and the routing paths configuration, we distinguish asynchronous PQP and synchronous PQP.

5.3.2. Asynchronous PQP

In this case, each two overlapping paths in the routing configuration have been assigned different wavelengths. Hence, each path can be assigned a number of wavelengths (one or more wavelengths) to send its LS bursts with loss-free transmission guarantee inside the OBS network. Whereas LS bursts can use only the wavelength(s) of their path, BE bursts are allowed to use any available wavelength at their OBS source node. The way of sending BE traffic (using the wavelength selection algorithm presented in Section 5.3.4) preserves statistical multiplexing and high resource utilization which represent the strengths of OBS networks. The way of sending LS traffic allows providing loss-free transmission inside the OBS network to this class of traffic. In case a LS burst contends with a BE burst, the LS burst is privileged and it can preempt the BE burst if necessary. Figure 42 shows the operation of PQP to schedule a LS burst both when the path to be used is asynchronous (asynchronous PQP) and when it is synchronized with overlapping paths (synchronous PQP).

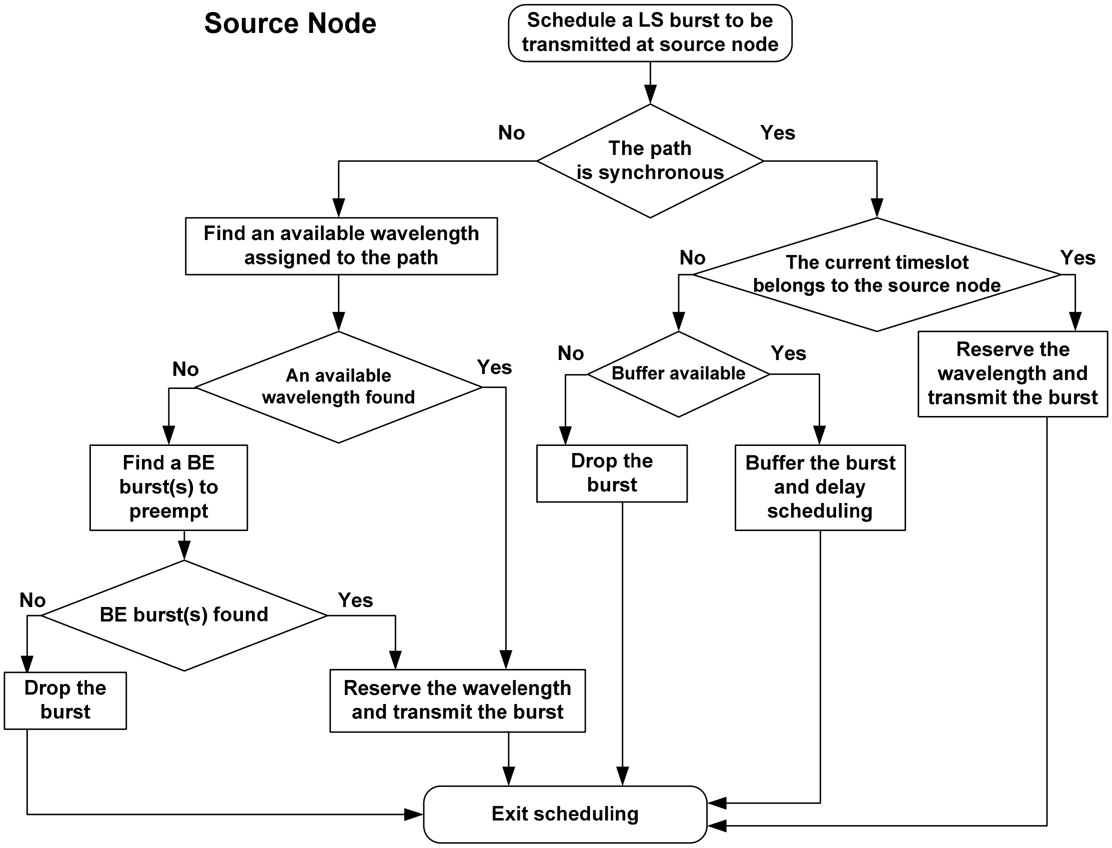


Figure 42. Operation of PQP (synchronous and asynchronous operation modes).

5.3.3. Synchronous PQP

In this case, some overlapping paths in the routing configuration (i.e., sharing at least one link) are assigned the same wavelength. Hence, to provide absolute QoS provisioning for LS bursts with loss-free guarantee inside the OBS network, a sub-wavelength utilization scheme is required. Thus, we propose a limited scale synchronization scheme, called Path-based Synchronous Transmission scheme (PST). PST allows overlapping paths to provide absolute QoS provisioning to LS traffic. The details of

PST are presented in Section 5.5. Once overlapping paths are synchronized, each path is allowed to carry LS traffic during a period of time called *timeslot*. More specifically, each path can only transmit an LS burst during its guaranteed timeslot. If the amount of arriving LS traffic is larger than the capacity of a path to carry (or to buffer) LS traffic, the exceeding LS traffic is dropped due to buffer overflow and the corresponding client networks are notified. It is worth noting that only overlapping paths are synchronized using PST; in fact, paths that do not suffer from overlaps do not need to be synchronized and continue to operate in the asynchronous mode. Meanwhile, like in the asynchronous case, BE bursts are allowed to use any available wavelength, asynchronously, at their OBS source nodes. Figure 42 shows the operation of synchronous PQP where some paths have a synchronous operation mode whereas the other paths continue to operate at the asynchronous operation mode.

5.3.4. Path-based best effort wavelength selection

Path-based Best Effort Wavelength Selection (PBEWS) is a wavelength selection scheme that aims to improve the performance of best effort traffic and to preserve statistical multiplexing gain and high resource utilization of OBS networks. PBEWS relies on the routing and wavelength assignment approach (OBSRWA, presented in Section 5.4) to select a wavelength on which a BE burst will be sent to its destination. To explain the operation of PBEWS, and without loss of generality, we suppose that each path $p_i, i \in \{0,1,.., N(N-1)-1\}$ (where N is the number of edge nodes in the OBS network) has been assigned a single wavelength $w_j, j \in \{0,1,...,W-1\}$ (where W is the number of wavelengths in each fiber link of the OBS network) by OBSRWA. To send a BE burst on path p_i, a source OBS node searches an available wavelength starting from wavelength w_j then w_{j+1} to w_{W-1} and then from wavelength w_0 to wavelength w_{j-1}. Figure 43 shows the order in which an available wavelength is searched to send a BE burst on path p_i. The rationale behind this technique to search available wavelengths is that each path searches available wavelengths in different order to reduce the probability that two overlapping paths choose the same wavelength for BE bursts and to increase BE traffic isolation among different paths, especially at low and moderate traffic loads; this is clearly not the case of classical wavelength selection algorithms (e.g., first-fit and random policies) which perform wavelength selection based on the local information in each node.

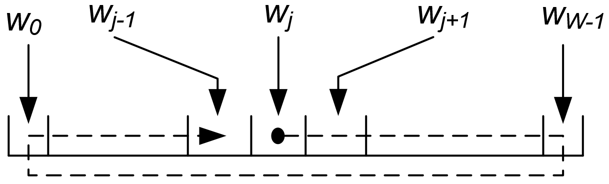


Figure 43. The order of searching an available wavelength for a BE burst on path p_i assigned a wavelength w_j.

It is worth noting that PBEWS can be considered as general wavelength selection scheme that can be also used in the classless case (i.e., without QoS provisioning).

5.4. Routing and wavelength assignment approach

In this Section, we present the routing paths optimization part of OBSRWA by presenting (1) the informal description and the aim of routing paths optimization; and (2) an exact and efficient formulation of the problem as an ILP model. Then, we present the wavelength assignment part of OBSRWA by presenting an ILP formulation to the problem and a tabu search algorithm to resolve large instances of this problem.

5.4.1. Routing paths optimization

5.4.1.1. Overview

The objective of routing paths optimization in OBSRWA is to find the optimal routing paths configuration that reduces the number of required wavelengths to establish completely non overlapping paths, or at least, minimizes the cases of overlaps between paths. This is equivalent to finding a routing paths configuration for which the chromatic number (the minimum number of colors required to color a graph properly) of the conflict graph is minimum. For example, let us consider the case shown in Figure 44. In this example, four paths (P_0 to P_3) share the fiber link from node A to node B; these paths form a clique (i.e., a complete subgraph) in the conflict graph of this routing paths configuration since they are all overlapping with one another. Hence, the number of wavelengths required to alleviate path overlapping in this case is the number of paths itself (i.e., four wavelengths).

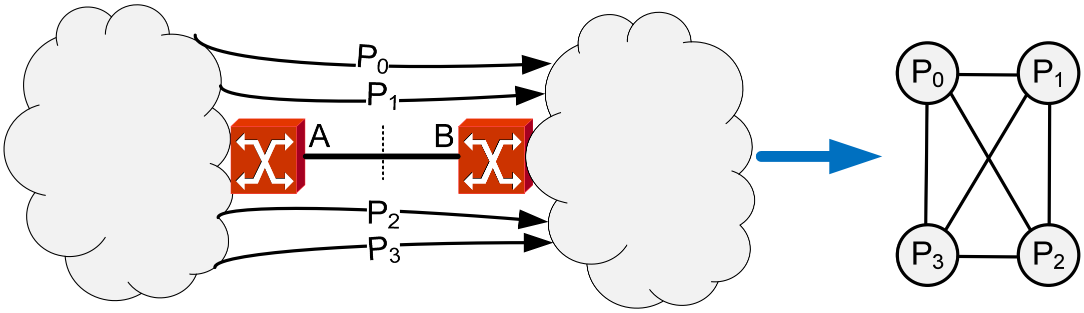


Figure 44. Example of four paths sharing the fiber link (A, B).

The size of the maximum clique in the conflict graph (i.e., a complete subgraph of maximum size) is a lower bound on the chromatic number of the conflict graph. However, finding the maximum clique size is an NP-Hard problem [108]. In the following, instead of trying to find a routing paths configuration that minimizes the size of the maximum clique in the conflict graph, we optimize a simpler quantity to handle, namely, the largest flow (number of paths) traversing any link (i.e., a lower bound on maximum clique size and the number of wavelengths needed for contention-free routing).

5.4.1.2. The exact formulation

We formulate the problem of finding routing paths between the edge nodes of the OBS network as an ILP model. We model the OBS network as a graph $G(V, E)$ where V

is the set of nodes n_i, $i \in \{0,1,...,N-1\}$ and E is the set of directed links. Each path (n_i, n_j) is denoted by an index $k \in \{0,1,...,M-1\}$ where M is equal to $N(N-1)$ and n_i is denoted by u_k and n_j is denoted by v_k.

ILP model A

Input:

$N(n_i)$ The set of neighbors of node n_i.

D The maximum authorized length of each path in the network.

Variables:

x_{ij}^k Binary variables which take value 1 if path k uses link (n_i, n_j); 0 otherwise.

Objective:

$$Minimize\left[\underset{(n_i,n_j)\in E}{Max}\sum_k x_{ij}^k\right] \tag{5.1}$$

Subject to:

$$\sum_{n_j \in N(n_i)} x_{ij}^k - \sum_{n_j \in N(n_i)} x_{ji}^k = \begin{cases} 1; & \text{if } n_i = u_k \\ -1; & \text{if } n_i = v_k \quad \forall k,i \\ 0; & \text{otherwise} \end{cases} \tag{5.2}$$

$$\sum_{(n_i,n_j)\in E} x_{ij}^k \le D \quad \forall k \tag{5.3}$$

$$x_{u_k v_k}^k = 1 \quad \forall (u_k, v_k) \in E \tag{5.4}$$

$$\sum_{n_j \in N(n_i)} x_{ij}^k \le 1 \quad \forall k,i \tag{5.5}$$

$$\sum_{n_j \in N(n_i)} x_{ji}^k \le 1 \quad \forall k,i \tag{5.6}$$

Bounds:

$x_{ij}^{k} = 0, 1 \quad i, j = 0,1,..., N-1 \quad k = 0,1,..., (M-1)$

To make the objective function (5.1) linear, we rewrite it as (5.1)' and add constraints (5.7):

$$\textit{Minimize } Y \tag{5.1)'$$

$$Y \geq \sum_{k=0}^{M-1} x_{ij}^{k} \quad \forall (n_i, n_j) \in E \tag{5.7}$$

Constraints (5.2) are classical flow conservation constraints. Constraints (5.3) limit the maximum length of routing paths in terms of the number of hops; parameter D in these constraints is specific to each network topology and it can be determined based on the diameter of the OBS network. Constraints (5.4) ensure that if two nodes are adjacent in the network (i.e., they are directly connected by a link), then they have to communicate directly through the link connecting them. This obviously optimizes the routing paths since one-hop communications, whenever they exist, are used instead of multi-hop communications. Constraints (5.5) and (5.6) guarantee that the produced routing paths are loopless, i.e., each node in the network is visited at most once by a routing path. The resolution of the above model determines a solution to the routing paths optimization problem; this solution will be contained in variables x_{ij}^{k}. This problem is a variant of the Constrained Shortest Path (CSP) problem which is NP-Hard [109]; however, computational experiments with respect to model A show that this model can be resolved efficiently for instances of the size encountered in real networks (e.g., SNDlib instances [106], related results are presented in Section 5.6.1); this obviates the need for using heuristic methods to resolve the problem when dealing with instances in that size range.

5.4.2. Wavelength assignment

5.4.2.1. Overview

The aim of our wavelength assignment scheme is to assign, as much as possible, different wavelengths to overlapping paths even when the number of available wavelengths is smaller than the number of wavelengths required to avoid completely overlaps between paths; this will allow providing QoS provisioning while reducing, or ideally, completely avoiding the need to synchronize overlapping paths.

Assigning wavelengths to routing paths while allowing, if necessary, conflicting assignments (i.e., assign the same wavelength to overleaping paths) is a characteristic specific to our scheme; indeed, existing wavelength assignment schemes aim, generally, to assign a distinct wavelength to each path/connection which is referred to as the wavelength-distinct constraint [1].

5.4.2.2. The required number of wavelengths

To assign wavelengths to the routing paths, we consider the routing paths configuration returned by the resolution of ILP model A; this is a routing paths configuration for which wavelength assignment can be performed using the minimum number of wavelengths. To obtain a quite accurate approximation of the required number of wavelengths to perform wavelength assignment, we construct the conflict graph of the routing paths configuration, denoted $G'(V',E')$, where V' is the set of all routing paths ($|V'| = N(N-1) = M$) and E' is the set of undirected links. A link l exists between two vertices in the graph G' if the paths corresponding to these vertices share at least one link in the routing configuration. The problem of assigning a wavelength to each path becomes the problem of coloring the vertices of graph G'; the question we need to answer is: what is the number of colors required to color G' properly (i.e., assign distinct colors to adjacent nodes in G')? This is, obviously, the ideal case where conflicting paths will be completely isolated using wavelength assignment. The maximum number of paths traversing a single link is a lower bound on the chromatic number of the conflict graph; hence, by comparing the cost of ILP model A solution (Eq. (5.1)') to the number of available wavelengths W, we can determine whether there will be overlapping paths after wavelength assignment (specifically, if W is less than the cost of ILP model A solution). Nevertheless, even if the number of available wavelengths W is smaller than the chromatic number of G', we can color G' while minimizing the cases where adjacent vertices have the same color (i.e., the cases of overlapping paths).

5.4.2.3. Exact formulation

We formulate the problem of assigning wavelengths to routing paths while minimizing the number of overlaps/conflicts between routing paths as an ILP model as follows.

ILP model B

Input:

$G'(V',E')$	The conflict graph where each vertex i in V' represents a path and a link (i,j) exists only if path i overlaps with path j in the routing paths configuration.
W	The number of available colors (wavelengths) to color graph $G'(V',E')$.

Variables:

x'^k_i	Binary variables which take value 1 if node i is colored with color k; 0 otherwise.
y_{ij}	Binary variables which take value 1 if nodes i and j such that $(i,j)\in E'$ are colored with the same color; 0 otherwise.

Objective:

$$Minimize \sum_{(i,j)\in E'} y_{ij} \qquad (5.8)$$

Subject to:

$$\sum_{k=0}^{M-1} x_i'^k = 1 \quad \forall i \in V' \tag{5.9}$$

$$x_i'^k + x_j'^k - y_{ij} \le 1 \quad \forall (i,j) \in E', \forall k \tag{5.10}$$

Bounds:

$x_i'^k = 0, 1 \quad y_{ij} = 0, 1 \quad (i,j) \in E' \ k = 0,1,\ldots,(M-1)$

The objective function (5.8) of the model aims at minimizing the cases where two adjacent vertices in the conflict graph (overlapping paths) are colored with the same color. Constraints (5.9) constrain each vertex to be colored with one and only one color. Constraints (5.10) make the link between variables $x_i'^k$ and variables y_{ij}, i.e., if variables $x_i'^k$ and $x_j'^k$ assume value 1, variable y_{ij} should assume value 1. When exactly one variable $x_i'^k$ or $x_j'^k$ assumes value 1, variable y_{ij} should assume value 0 since we have a minimization problem.

This problem is a variant of the Vertex Coloring Problem (VCP), which is known to be NP-hard [110]. In addition, computational experiments with respect to model B (Section 5.6.1) confirm the difficulty to resolve exactly large instances of this problem. Hence, we propose a tabu search algorithm to deal with large instances of this problem.

5.4.2.4. Tabu search algorithm

The proposed tabu search algorithm aims to find, efficiently, near-optimal solutions to the wavelength assignment problem in a reasonable time. Tabu search has been proposed by Glover in 1986 [102]. This meta-heuristic searches the best feasible solution starting from the neighbourhood of an initial solution. The process is repeated until a maximum number of iterations without improvement is reached. Tabu search avoids cycling while searching solutions by forbidding moves that lead from the current solution to a previously visited solution; these moves are called tabu moves and they are stored in a list, called *tabu list*. In our tabu search algorithm, a move consists of changing the color of a vertex in the conflict graph. Also, we use a variable size tabu list; i.e., each new entry in the tabu list has to be stored for a random number of iterations of the algorithm. Algorithm 4 shows the pseudo code of the proposed tabu search algorithm.

Algorithm 4. Tabu search algorithm for wavelength assignment.

```
Begin
 Step 0: Initialization
    Find an initial solution $P_0 P_0$ using the greedy heuristic in Algorithm 5;
    Best_Solution = $P_0$;
    Best_Cost = Cost ($P_0$);
    Iteration_Number = 0;
 Step 1: Loop
  While(Iteration_Number < Max_Iterations)
     K = Iteration_Number;
     Find in the neighbourhood of current solution $P_K$ the neighbour corresponding to the best solution $P_{K+1}$ $P_k$obtained by a non-tabu move, i.e., by changing the color of a given node $V_i$ From color $C_l$ to color $C_m$, choose the solution which minimizes the maximum number of conflicts per wavelength to break ties;
     Add the pair ($V_i$, $C_l$) to the tabu list T;
     If (Cost ($P_{k+1}$) < Best_Cost) then
        Best_Solution = $P_{k+1}$;
        Best_Cost = Cost($P_{k+1} P_{k+1}$);
        Iteration_Number = 0;
     End If
     Iteration_Number++;
  End While
End
```

Algorithm 5. Initialization greedy heuristic for wavelength assignment.

```
Begin
 Step 0: Initialization
   $colored = \{\}$; /* the set of colored nodes */
   $uncolored = V'$; /* the set of uncolored nodes */
   $colors = \{0,1,...,W-1\}$; /* the set of available colors */
   Var $i$;
 Step 1: Loop
     While ($\overline{S} \neq \varnothing$)
       $i$ = the node with the highest degree in $uncolored$;
       Assign to $i$ the least used color in its neighborhood (choose a random color to break ties);
       $colored = colored \bigcup \{i\}$;
       $uncolored = uncolored - \{i\}$;
     End While
End
```

The algorithm begins by the initialization step where an initial feasible solution has to be found; we use the greedy heuristic presented in Algorithm 5 to find this initial solution. The initialization heuristic starts by initializing the sets: *colored* vertices, *uncolored* vertices and *colors*. After that, at each iteration, the uncolored vertex with the highest degree is selected and the least used color in the neighborhood of this vertex is assigned to it (ideally, an unused color); ties are broken by selecting a random color. At the end of each iteration, *colored* and *uncolored* sets are updated. The output of this algorithm is the conflict graph $G'(V',E')$ with all of its vertices *colored*. After the initialization step of the tabu search algorithm, at each iteration of the loop, the best solution P_{k+1} in the neighborhood of the current solution P_k is found using a non-tabu move; a move consists of changing the color of a vertex in the conflict graph $G'(V',E')$ from a color C_l to another color C_m. Ties among solutions are broken by choosing the solution which minimizes the maximum number of conflicts per wavelength; this will allow producing solutions where conflicts are fairly distributed amongst the wavelengths and thus yielding smaller sets of paths to be synchronized. Tabu moves are stored in the tabu list, hence, the pair (V_i, C_l) representing the last move is added to the tabu list. After finding solution P_{k+1} the tabu search algorithm compares its cost to the cost of the best solution found so far and makes the necessary updates. This process continues until a maximum number of iterations without improving the known best cost is reached. The computational complexity of the loop in this algorithm is $O(N^2W)$.

5.5. Path-based synchronous transmission scheme

5.5.1. Overview

In the case when the routing and wavelength assignment approach yields a routing paths configuration where some paths overlap, we propose to use synchronization to provide absolute QoS provisioning at sub-wavelength level; to the best of our knowledge, this is the first time that a hybrid OBS architecture is proposed where synchronous paths and asynchronous paths are operating together and where only subsets of paths are synchronized with each other. The proposed scheme is called, Path-based Synchronous Transmission scheme (PST).

The basic idea behind PST is to synchronize the transmissions of overlapping paths, i.e., paths using the same wavelength on the same link. However, to synchronize overlapping paths, each path has to transmit a limited amount of guaranteed traffic (a sub-wavelength amount); otherwise, contentions will continue to occur. We call this requirement *capacity constraint*. Our objective is to maximize the amount of guaranteed traffic that can be carried by each path. In the following, we propose an ILP formulation to the capacity constraint and present the synchronization protocol.

5.5.2. Capacity constraint problem

We consider again the conflict graph G' where vertices (corresponding to routing paths) are colored with wavelengths $\lambda_0, \lambda_2, ..., \lambda_{W-1}$. Also, we consider conflict sub-graphs $G'_{\lambda_0}(V'_0, E'_0)$, $G'_{\lambda_1}(V'_1, E'_1)$, ..., $G'_{\lambda_{W-1}}(V'_{w-1}, E'_{w-1})$ where each sub-graph $G'_{\lambda_i}(V'_i, E'_i)$ is composed of the set of vertices $V'_i \subseteq V'$ colored with wavelength λ_i and $E'_i \subseteq E'$ is the set of edges connecting two vertices colored with λ_i. Each conflict sub-graph $G'_{\lambda_i}(V'_i, E'_i)$ is composed of a set of *connected components* $\{G'^{j}_{\lambda_i}(V'^{j}_i, E'^{j}_i) \mid j = 0,1,...,L\}$ where each connected component $G'^{j}_{\lambda_i}$ represents a connected graph in $G'_{\lambda_i}(V'_i, E'_i)$. Figure 45 shows an example of eight paths operating on the same wavelength (wavelength 2) and forming three connected components in the conflict sub-graph of wavelength 2.

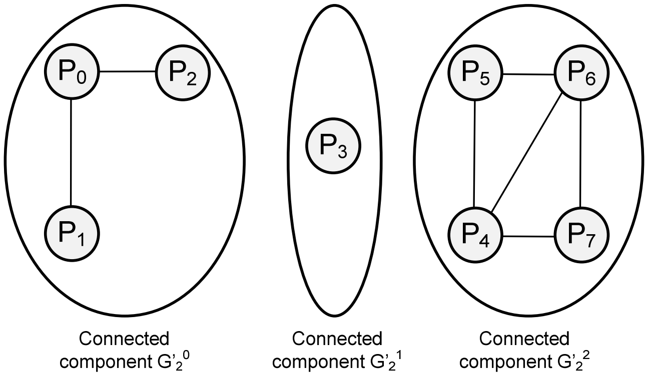


Figure 45. Example of three connected components of paths assigned wavelength 2.

For ease of reading, in the rest of this section, we call a vertex in the conflict sub-graph a path.

The minimum capacity of each path k belonging to the connected component $G'^{j}_{\lambda_i}$ to carry guaranteed traffic is given by:

$$C^k_{\min} = \frac{C_{\lambda_i}}{\chi(G'^{j}_{\lambda_i})} \tag{5.11}$$

and the maximum capacity of each path k belonging to the connected component $G'^{j}_{\lambda_i}$ to carry guaranteed traffic is given by:

$$\begin{aligned} C^{k}_{\max} &= \frac{C_{\lambda_i}}{\chi(G'^{j}_{\lambda_i})}(\chi(G'^{j}_{\lambda_i}) - d_{\chi}(k)) \\ &= C_{\lambda_i} - \frac{C_{\lambda_i} d_{\chi}(k)}{\chi(G'^{j}_{\lambda_i})} \end{aligned} \tag{5.12}$$

where C_{λ_i} is the bandwidth capacity of wavelength λ_i, $\chi(G'^{j}_{\lambda_i})$ is the chromatic number of the connected component $G'^{j}_{\lambda_i}$ (i.e., the minimum number of colors required for a proper coloring of $G'^{j}_{\lambda_i}$) and $d_{\chi}(k)$ is the *chromatic degree* of node k (i.e., the number of colors required to color its neighbors in a proper coloring of $G'^{j}_{\lambda_i}$). Eq. (5.11) and Eq. (5.12) determine the minimum and the maximum capacities of each path to carry guaranteed traffic, respectively. The objective is to guarantee that the amount of guaranteed traffic on each wavelength and each link does never exceed the bandwidth capacity of the wavelength on the link when all the paths are transmitting guaranteed traffic simultaneously. Whereas each path will have at least the minimum capacity (Eq. (5.11)), some paths could have an additional capacity (up to the maximum capacity in Eq. (5.12)); this is because the number of paths in the neighbourhood of a path (including itself) could be smaller than the chromatic number of its connected component $\chi(G'^{j}_{\lambda_i})$.

For better understanding, let us consider the connected component conflict sub-graph composed of 6 nodes and 9 edges in Figure 46. In this sub-graph, the chromatic number is 4, the chromatic degrees of paths *a*, *b*, *e* and *f* are all 3 and the chromatic degrees of paths *c* and *d* are all 1. Thus, the minimum capacity of all of the paths is $\frac{1}{4}$ the capacity of the wavelength; this is also the maximum capacity of paths *a*, *b*, *e* and *f*; the maximum capacity of paths *c* and *d* is $\frac{3}{4}$ the capacity of the wavelength. Since paths *c* and *d* are neighbors in the connected component, the difference between the maximum capacity and the minimum capacity of each one of them (i.e., $\frac{3}{4}-\frac{1}{4}=\frac{2}{4}$) is shared between the two paths. Hence, a good solution will allocate an additional $\frac{1}{4}$ the wavelength capacity to each path.

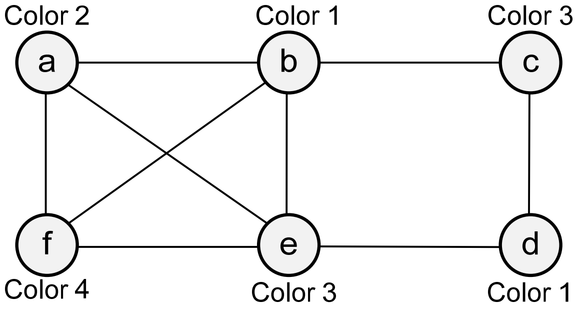


Figure 46. Example of a connected component conflict sub-graph of paths operating on the same wavelength.

To satisfy the capacity constraint while synchronizing guaranteed traffic transmission on each path k of each connected component $G'^{j}_{\lambda_i}$, we define a fixed timeslot duration τ corresponding to a fixed burst duration; τ is used to express each path capacity in terms of timeslots rather than a raw capacity (e.g., in Gbps). Thus, we can express the capacity constraint in Eq. (5.11) and Eq. (5.12) for a connected component $G'^{j}_{\lambda_i}$ as follows: each path $k \in V'^{j}_{i}$ is able to use a number of timeslots (at least one) to send guaranteed traffic over a period of time, called *a cycle of timeslots*, equals to $\chi(G'^{j}_{\lambda_i}) \cdot \tau$ (i.e., $\chi(G'^{j}_{\lambda_i})$ timeslots). In the example shown in Figure 46, paths *a*, *b*, *e* and *f* can use at most 1 timeslot in a cycle of 4 timeslots while paths *c* and *d* can use at most 3 timeslots in a cycle of 4 timeslots.

The problem of determining the capacity of each path in $G'^{j}_{\lambda_i}$ can be formulated using two ILP models as follows. The first ILP model (ILP model C) determines the maximum number of timeslots that can be used by the set of paths in the connected component, denoted $\aleph$. The second ILP model (ILP model D) determines the set of timeslots of each path. ILP model C is defined as follows.

ILP model C

Input:

$G'^{j}_{\lambda_i}$ The j^{th} connected component conflict sub-graph on wavelength λ_i.

$\chi(G'^{j}_{\lambda_i})$ The chromatic number of $G'^{j}_{\lambda_i}$.

Variables:

t_{kh} Binary variables which take value 1 if timeslot h is assigned to path k; 0 otherwise.

Objective:

$$Maximize \quad \aleph = \sum_{k \in V'^{j}_{i}} \sum_{h=1}^{\chi(G'^{j}_{\lambda_i})} t_{kh} \tag{5.13}$$

Subject to:

$$\sum_{h=1}^{\chi(G'^{j}_{\lambda_i})} t_{kh} \geq 1 \quad \forall k \tag{5.14}$$

$$t_{kh} + t_{mh} \leq 1 \quad \forall h, \forall (k,m) \in E'^{j}_{i} \tag{5.15}$$

Bounds:

$t_{kh} = 0, 1 \quad k, m = 1,2,...,|V'^{j}_{i}| \quad h = 1,2,...,\chi(G'^{j}_{\lambda_i})$

The objective function (5.13) maximizes the total number of timeslots allocated to the paths of the connected component $G'^{j}_{\lambda_i}$. Constraints (5.14) state that each path should be allocated at least one timeslot and constraints (5.15) state that two adjacent paths in the connected component should not be allocated the same timeslot. The solution of model C is the maximum number of timeslots that can be allocated to paths in the connected component $G'^{j}_{\lambda_i}$. This value is used in ILP model D to determine the optimal allocation of timeslots to paths in $G'^{j}_{\lambda_i}$ as follows.

ILP mode D

Input:

$G'^{j}_{\lambda_i}$	The j^{th} connected component conflict sub-graph on wavelength λ_i.
$\chi(G'^{j}_{\lambda_i})$	The chromatic number of $G'^{j}_{\lambda_i}$.
$\aleph$	The maximum number of timeslots that can be allocated to paths in $G'^{j}_{\lambda_i}$.

Variables:

t_{kh}	Binary variables which take value 1 if timeslot h is assigned to path k; 0 otherwise.
s_{kl}	Binary variables which take value 1 if path k uses *at least* l timeslots; 0 otherwise.
$\vartheta : \mathrm{N} \rightarrow \mathrm{N}$	A function where $\vartheta(l)$ determines the gain when assigning an additional l^{th} timeslot to any path in $G'^{j}_{\lambda_i}$ and where $\vartheta(2) > \vartheta(3) > ... > \vartheta(\chi(G'^{j}_{\lambda_i}))$ and $\vartheta(l) = 2^{\chi(G'^{j}_{\lambda_i})-l}$.

Objective:

$$Maximize \sum_{l=2}^{\chi(G'^{j}_{\lambda_i})} [\vartheta(l) \sum_{k \in v'_i} s_{kl}] \qquad (5.16)$$

Subject to:

$$\sum_{h=1}^{\chi(G'^{j}_{\lambda_i})} t_{kh} - \sum_{l=1}^{\chi(G'^{j}_{\lambda_i})} s_{kl} = 0 \quad \forall k \qquad (5.17)$$

$$s_{k1} = 1 \quad \forall k \qquad (5.18)$$

$$t_{kh} + t_{mh} \le 1 \quad \forall (k,m) \in E_i'^{j}, \quad \forall h \in 1..\chi(G_{\lambda_i}'^{j}) \tag{5.19}$$

$$\sum_{k \in V_i'^{j}} \sum_{l=1}^{\chi(G_{\lambda_i}'^{j})} s_{kl} = \aleph \tag{5.20}$$

Bounds:

$$t_{kh} = 0,1 \quad s_{kl} = 0,1 \quad k,m = 1,2,...,|V_i'^{j}| \quad h,l = 1,2,...,\chi(G_{\lambda_i}'^{j})$$

The objective function (5.16) maximizes the total number of allocated timeslots to the paths while guaranteeing that timeslots will be fairly allocated; fairness is realized using function ϑ whose values are decreasing power of 2; hence, if, for example, two adjacent paths in $G_{\lambda_i}'^{j}$ could both be allocated an additional timeslot and one of them has already been allocated two timeslots while the other one has already been allocated only one timeslot, function ϑ makes it more valuable to allocate the additional timeslot to the path which has only one timeslot. Constraints (5.17) establish the relation between variables t_{kh} and variables s_{kl}. Constraints (5.18) state that each path should be allocated at least one timeslot. Constraints (5.19) state that two adjacent paths cannot be assigned the same timeslot. Constraints (5.20) state that the total number of allocated timeslots in the connected component $G_{\lambda_i}'^{j}$ should be equal to $\aleph$.

The capacity constraint problem formulated above is a variant of Vertex Multi-Coloring Problem (VMCP) which is known to be NP-hard [110]; however, computational experiments in Section 5.6.1 show that our formulation is computationally tractable for instances of the size encountered in real networks (e.g., SNDlib instances [106]).

5.5.3. The synchronization protocol

To synchronize the transmissions of overlapping paths, each source node in the OBS network uses a timer for each path to determine the beginning of timeslots. The problem is that, in the absence of a global clocking system, timers of different nodes are, generally, not synchronized. To overcome this problem, we propose a synchronization scheme, called Tree-based Path Synchronization scheme (TPS). The basic idea behind TPS is that each path k (represented by a vertex in a connected component conflict sub-graph $G'^{j}_{\lambda_i}$) *synchronizes its transmission with only one other path* l in $G'^{j}_{\lambda_i}$ where k and l are overlapping paths (i.e., neighbors in $G'^{j}_{\lambda_i}$). l is called the *parent* of k and k is a called a *child* of l. A given path could have at most one parent but could have many children. It is clear that the dependencies between paths operating on the same connected component conflict sub-graph form a tree which is rooted at a single path; we call this tree *synchronization tree*. The synchronization tree makes all of its paths synchronized (directly or indirectly) to the single path at its root. We propose that the synchronization tree of a connected component $G'^{j}_{\lambda_i}$, denoted $S(G'^{j}_{\lambda_i})$, be the minimum spanning tree of $G'^{j}_{\lambda_i}$ rooted at the vertex with the biggest nodal degree in $G'^{j}_{\lambda_i}$; the proposed spanning tree will minimize the depth of the tree $S(G'^{j}_{\lambda_i})$ which means that the synchronization process will take less time to converge. $S(G'^{j}_{\lambda_i})$ is then searched (e.g., in breadth-first order) and each path is attributed a *shift value* which represents the shift between the lowest timeslot of the current path and the lowest timeslot of its parent in $S(G'^{j}_{\lambda_i})$ (timeslots are ordered from 1 to $\chi(G'^{j}_{\lambda_i})$); the shift value could be a positive or a negative integer. Since the path at the root of $S(G'^{j}_{\lambda_i})$ has no parent path, its shift value is set to 0. Figure 47 shows the synchronization tree of the example shown in Figure 46 where, for each path, the lowest timeslot is shown at the left side and the shift value is shown at the right side.

The synchronization tree is rooted at path *b* which has the lowest timeslot 0 with a shift value set to 0; paths *a*, *c*, *e* and *f* are children of path b, have lowest timeslots 1, 2, 2 and 3 and have shift values -1, -2, -2 and -3, respectively; path *d* which is child of path *e* has lowest timeslot 0 and shift value 2. It is worth noting that since neighboring paths in $G'^{j}_{\lambda_i}$ have always different timeslots, a path and its parent in $S(G'^{j}_{\lambda_i})$ will never have the same timeslots, and hence, shift values are always different from 0, except for the path at the root of $S(G'^{j}_{\lambda_i})$.

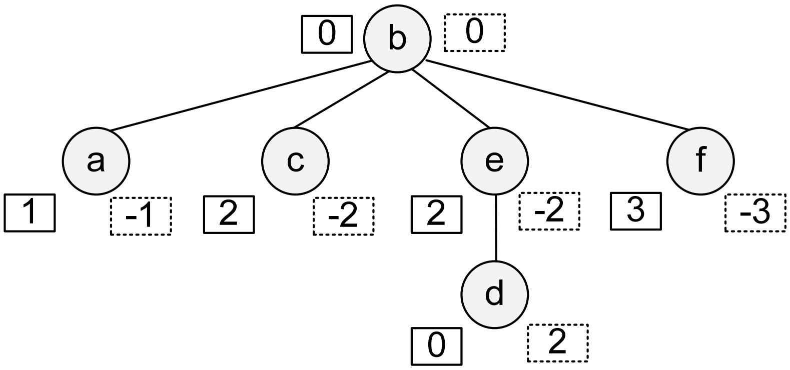


Figure 47. Example of a synchronization tree.

The synchronization schedule obtained from the synchronization tree is applied at the start of the OBS network operation; special packets, called *Sync* packets, are sent on control wavelength(s) where each path sends *Sync* packets corresponding to its lowest timeslot per a cycle of timeslots. The main fields of a *Sync* packet are presented in Figure 48 where:

- *Id*: the identifier of the packet;
- *Label*: the label of the sending path;
- *PLabel*: the label of the parent of the sending path;
- *Syn*: a Boolean field that assumes the value 1 if the sending path is already synchronized with its parent path; at the beginning of the synchronization phase, only the path at the root of the synchronization tree has this field set to 1;
- *PSyn*: a Boolean field indicating whether or not the parent path is synchronized;

- *ArrivalTime*: the arrival time of the corresponding timeslot (burst). Bursts are not sent before the path is synchronized; *ArrivalTime* is used to take into consideration the offset time which separates the control packet from its timeslot (burst);

- *PArrivalTime*: the arrival time of the parent path.

`Id`	**`Label`**	**`PLabel`**	**`Syn`**	**`PSyn`**	**`ArrivalTime`**	**`PArrivalTime`**

Figure 48. The main fields of Sync packet.

Whereas fields *Id*, *Label*, *PLabel*, *Syn* and *ArrivalTime* are set at the source node of the corresponding path (*ArrivalTime* is updated at each intermediate OBS node), *PSyn* and *PArrivalTime* are set at the first intermediate OBS node where the corresponding path and its parent overlap. To this end, intermediate OBS nodes keep track of the most recent information on *Syn* and *ArrivalTime* fields of different paths; this is realized in a lookup table where each entry corresponds to the label of a path. Upon receipt of Sync packet where field Syn is set to 0, the first intermediate OBS node where a path sending Sync packet and its parent path overlap: (a) drops Sync packet if the path and its parent path are not synchronized; (b) drops Sync packet and responds by a *ReSync* packet to the source node of the path if the path is not synchronized but its parent path is already synchronized; and (c) forwards Sync packet towards the destination of the path in order to synchronize potential children paths if the corresponding path is already synchronized. ReSync packet contains a numerical value representing the difference between the value of *ArrivalTime* field and the value of *PArrivalTime* filed. When a source OBS node receives a ReSync packet it compares the value in ReSync packet and its shift value to re-adjust its timers and transmission times, and hence, synchronizes its transmission with its parent path. Note that Sync packets with the field *Syn* set to 1 (i.e., the corresponding path is synchronized) are never dropped since they are required to synchronize the children of the corresponding path. Also, once a path is synchronized, it can start carrying real data and piggyback the content of sync packets with its control packets. After a period of time of the synchronization of a path, it can send (or piggyback) Sync packets less frequently in order to reduce the overhead on the control plane while maintaining the synchronization of overlapping paths.

The time needed for the synchronization process to converge and the required number of Sync packets is bounded by the depth of synchronization tree, i.e., O($depth(S(G'^{j}_{\lambda_i}))$).

5.5.4. Synchronization issues

One of the problems of synchronization is small scale fluctuations that may occur in the propagation delay of a wavelength on a link because of the environmental conditions (e.g., temperature) [15]. PST resolves this problem by adding a *guard time* between consecutive timeslots. Such a guard time will not necessarily reduce the throughput of the link since it is useful to perform some operations such as the set-up of switching fabrics. Also, even if the guard time is not sufficient to overcome the propagation delay fluctuations problem, a re-synchronization can be triggered at the time when an incorrect reception of a timeslot event (and hence, a contention between two timeslots) occurs. Fortunately, the re-synchronization scope will be the sub-tree of the synchronization tree rooted at the path for which an incorrect reception has been detected.

5.6. Numerical results

In this Section, we present numerical results of the resolution of ILP models (A, B, C and D) using CPLEX 10.11 solver [98] and we compare the results of ILP model B and the results of the proposed tabu search algorithm. Also, we present simulation results that show the performance of both asynchronous PQP and synchronous PQP. We use ns-2 simulator [93] and modules that implement OBS in ns-2 [94]. We use NSFNET topology with 14 nodes (Figure 49).

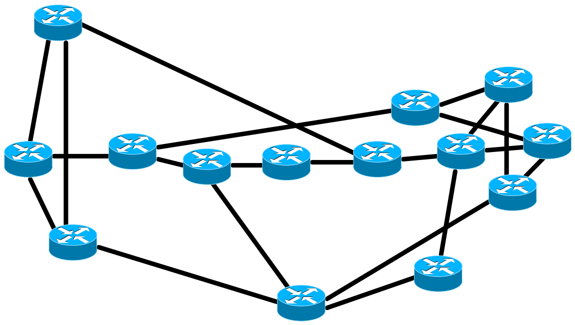


Figure 49. NSFNET topology.

We assume that each single fiber link is bidirectional and all links have the same number of wavelengths each one operating at 1Gbps. Since the number of required wavelengths to establish completely non-overlapping paths in NSFNET topology with OBSRWA solution is 13, we use 64 wavelengths for asynchronous PQP and 8 wavelengths for synchronous PQP; the timeslot duration is 0.01 s, i.e., the size of each burst is 1.25 MB (other values are simulated too); and the buffering capacity in each source node for each path is equal to the duration of one timeslot, unless stated otherwise. Each node in the network can generate, route and receive traffic (i.e., each node in the network is an edge and core node at the same time). Sources and destinations of traffic connections are generated randomly between any two nodes in the network, i.e., the traffic is dynamic and uniformly distributed over source nodes. The traffic load is expressed as the average traffic load per link. The capacity of a link is the sum of the capacities of all the wavelengths in this link. We use exponential ON/OFF traffic. We consider loss probability which is the main performance metric in buffer-less OBS networks and delay at the OBS source node for synchronous PQP. We call guaranteed traffic *Loss Sensitive (LS) traffic* and non-guaranteed traffic *Best Effort (BE) traffic.*

All the following results have a confidence level of 95%.

5.6.1. Results of ILP models (A, B, C and D) and tabu search algorithm

Table 3 displays numerical results of the exact resolution of the routing paths optimization ILP model (ILP model A) using CPLEX 10.11 solver [98] on twelve of the SNDlib instances benchmark [106]. For each instance, we report the number of nodes, the number of links, the objective function value (Eq. (5.1)') for Shortest Paths configuration Y(SP), the number of required wavelengths for SP configuration (the chromatic number of the conflict graph), the objective function value (Eq. (5.1)') for ILP model A solution Y*, the number of required wavelengths for ILP model A solution (the chromatic number of the conflict graph), the resolution time of ILP model A and the improvement in terms of required number of wavelengths by ILP model A solution (defined as [(number of required wavelengths for SP – number of required wavelengths for ILP model A solution) / number of required wavelengths for SP]. The number of required wavelengths is computed by determining the chromatic number of the conflict graph of the routing paths configuration. In our case, we use the well known DSATUR algorithm [111]. Results show that ILP model A reduces effectively the number of wavelengths required to establish non-overlapping paths. In fact, the improvement brought by ILP model A is between 13.33% (2 wavelengths) for the NSFNET instance and 63.27% (274 wavelengths) for ta2 instance; this will allow for establishing more non-overlapping paths between OBS edge nodes in order to provide QoS provisioning without the need for synchronization. Also, these results show that our approach to reduce the maximum number of paths traversing a single link in order to reduce the required number of wavelengths is correct. Indeed, the value of Y* and the number of required wavelengths are quite close for all of the instances.

With respect to ILP model A resolution, the response time is very reasonable. In fact, the biggest value is 59871.8 s (less than 17 hours) for germany50 instance which is acceptable for a design problem like in our case.

Table 3. Results of routing paths optimization ILP model (A).

Instance	# nodes	# links	Y(SP)	Required wavelengths for SP	Y^*	Required wavelengths for ILP model A solution	Resolution time (s)	Improvement
NSFNET (nobel-us)	14	21	15	15	13	13	0.73	13.33%
Atlanta	15	22	26	26	19	20	1	23.07 %
Newyork	16	49	12	13	8	8	0.85	38.64 %
France	25	45	51	51	34	36	8.9	29.41 %
janos-us	26	84	88	88	42	45	35.97	48.86 %
cost266	37	57	162	162	86	91	616.82	43.82 %
giul39	39	172	77	77	39	44	52428.3	42.85 %
janos-us-ca	39	122	184	184	108	108	777.45	41.30 %
pioro40	40	89	144	144	75	77	119.62	46.52 %
germany50	50	88	238	236	91	98	59871.8	58.47 %
zib54	54	81	336	336	147	148	520.99	55.95 %
ta2	65	108	433	433	151	159	3668.57	63.27 %

Table 4 shows results of wavelength assignment ILP model (ILP model B) and Tabu Search (TS) algorithm. We use NSFNET instance and vary the number of wavelengths from 6 wavelengths to 13 wavelengths. For each instance, we present the cost of the solution returned by CPLEX solver after a running time of at most 86400 s (i.e., 24 h), the status of the solution returned by CPLEX solver (i.e., Optimal or Feasible), the running time of CPLEX solver (useful if the optimal solution is found before 86400 s), the cost of TS algorithm solution (returned after 10000 iterations without improvement) and the running time of TS algorithm. Recall that the number of required wavelengths to assign one wavelength to each path in NSFNET topology without overlapping is 13 wavelengths. Hence, when the number of available wavelengths is 13, the optimal cost defined in Eq. (5.8) is 0; we observe that in this case both the exact resolution using CPLEX solver and TS algorithm find the optimal solution, however, TS algorithm is faster than CPLEX since it finds the optimal solution in 14.48 s versus 72.54 s for CPLEX. When the number of available wavelengths is smaller than 13, TS algorithm outperforms CPLEX in terms of the cost of the solution and the resolution time. For example, with 6 wavelengths, the cost of CPLEX solution is 328 obtained after 86400 s of running time; the cost of TS algorithm solution is 198, obtained after 4888.35 s of running time; this is roughly 39% improvement in the cost of the solution with a shorter response time.

Table 4. Results of wavelength assignment ILP model (B) and Tabu Search (TS) algorithm

Number of wavelengths	**ILP model B solution cost**	**ILP model B solution status**	**ILP model B CPU time (s)**	**TS solution cost**	**TS CPU time (s)**
6	328	Feasible	86400	194	5615.46
8	152	Feasible	86400	94	3593.40
10	58	Feasible	86400	40	1038.61
12	10	Feasible	86400	8	107.23
13	0	Optimal	72.54	0	14.48

Table 5 shows results of the exact resolution of ILP model C and ILP model D.

Table 5. Results of ILP model C and ILP model D for Janos-us-ca instance when varying the number of wavelengths

Available wavelengths	**Number of nodes in max. C.C.**	**Number of links in max. C.C.**	**Number of timeslots in a cycle**	**ILP model C cost (timeslots)**	**ILP model C resolution time (s)**	**ILP model D resolution time (s)**
6	241	7906	21	605	787.43	17793.6
8	172	3502	17	429	34.49	54.88
16	82	612	8	167	0.68	1.98
32	25	82	4	37	0	0.02
64	6	10	2	6	0	0

We use *janos-us-ca* instance (39 nodes and 122 links) and vary the number of wavelengths from 6 to 64 wavelengths. Each row in the table presents the number of available wavelengths, the number of nodes in the maximum Connected Component (C.C.), the number of links in the maximum C.C., the chromatic number of the maximum C.C. (i.e., the number of timeslots in a cycle of timeslots), the cost of the solution returned by CPLEX for ILP model C; this cost represents the capacity of the maximum C.C. in terms of the number of timeslots, the resolution time of ILP model C and the resolution time of ILP model D. We observe that even large instances could be resolved exactly in a reasonable time; for example, when the number of available wavelengths is 6, we have a maximum connected component of 241 nodes and 7906 links; this instance is resolved in 787.43 s and 17793.6 s for ILP model C and ILP model D, respectively. This proves that ILP models C and D are efficient even for large instances of the timeslot allocation problem.

5.6.2. Results of asynchronous PQP

In this Section, we consider the performance of asynchronous PQP (i.e., when the number of available wavelengths allows establishing non-overlapping paths between each pair of edge nodes). We compare the performance of PQP to the performance of AFQD [107] which is an absolute QoS provisioning scheme for OBS networks that aims to guarantee loss free transmission inside the OBS network for LS traffic by assigning a number of wavelengths to each OBS edge node; hence, the ratio to measure the amount of guaranteed traffic in AFQD is 1/N (where N is the number of edge nodes in the OBS network). Since LS bursts are sent with loss-free guarantee inside the OBS networks, we report their blocking probability at the access of the OBS network (i.e., admission control blocking probability).

Figure 50 shows the overall loss probability of PQP and AFQD when the ratio of LS traffic is 1/N (1/14 for NSFNET topology) and the number of wavelengths is 64. We observe that PQP outperforms significantly AFQD. In fact, whereas the mean loss probability of AFQD (over all the loads) is 0.29, the mean loss probability of PQP is 0.17. In addition, the loss probability of PQP at load 0.1 is 4x10-6.

Figure 51 shows the blocking probability of LS traffic when varying its amount (1/N and 2/N). We observe that at the same amount of LS traffic, PQP decreases the blocking of LS traffic by orders of magnitude. Indeed, when the amount of LS traffic is 1/N, the mean blocking probability of PQP (over all the loads) is 1.4x10-4 and the mean blocking probability of AFQD is 4.3x10-2. We conclude that PQP is able to accommodate more LS traffic inside the OBS network with loss-free transmission guarantee.

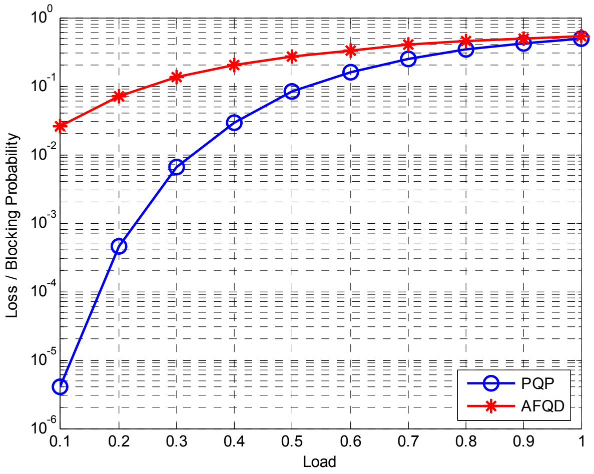


Figure 50. Overall loss probability vs. load for PQP and AFQD.

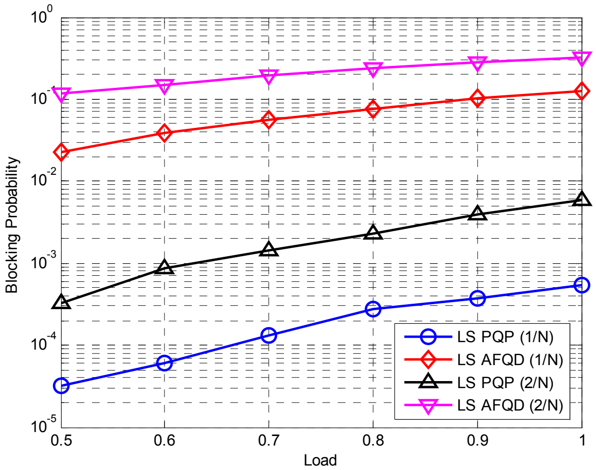


Figure 51. Blocking probability for LS traffic.

5.6.3. Results of synchronous PQP

We compare the performance of synchronous PQP to TSA scheme (see [14] in 5.2) and we report packet blocking probability outside the OBS network due to buffer overflow and the delay incurred by delaying bursts while waiting for a guaranteed timeslot.

Since both synchronous PQP and TSA offer loss-free transmission inside the OBS network for guaranteed traffic using synchronization, we compare PQP and TSA in terms of the number of required wavelengths. To do so, we consider the scenario reported in [14] (Fig. 10 in [14]) where NSFNET topology is used, the bandwidth of each wavelength is 2.5 Gbps and each OBS edge node sends an amount of traffic of 0.25 Gbps on each path (i.e., to each OBS edge node). Hence, each node generates an overall amount of traffic of 3.25 Gbps. Figure 52 shows the number of required wavelengths for TSA before applying a technique, called multiplexing optimization, which aims to reduce the number of required wavelengths; in this case, the number of required wavelengths is 182 wavelengths; after applying multiplexing optimization, the number of required wavelengths is 68 wavelengths. The number of required wavelengths for PQP in this scenario is as low as 8 wavelengths to guarantee zero blocking at the access of the OBS network and loss-free transmission inside the OBS network; this is roughly 95% reduction compared to the case of TSA without multiplexing optimization and 88% reduction compared to the case of TSA with multiplexing optimization. This shows clearly that PQP optimizes the required number of wavelengths and the bandwidth utilization while guaranteeing loss free transmission for guaranteed traffic.

Figure 53 shows the QoS differentiation capability of synchronous PQP in terms of loss probability when the number of wavelengths is 8. The proportion of LS traffic is 50% of the overall traffic and the proportion of BE traffic constitutes the rest of the traffic. Whereas the mean loss probability of BE traffic (over all the loads) is 0.32, the blocking probability of LS traffic is as low as 0.011. In addition, at traffic loads smaller than 0.7, the blocking probability of LS traffic is less than 1%. This shows clearly the capability of PQP to provide QoS differentiation among high priority and low priority traffic.

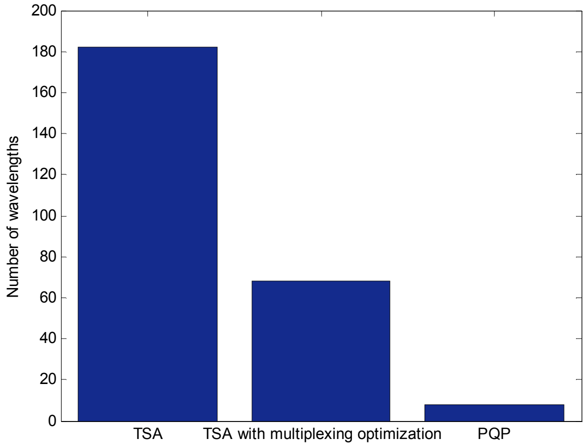


Figure 52. Comparison of PQP and TSA

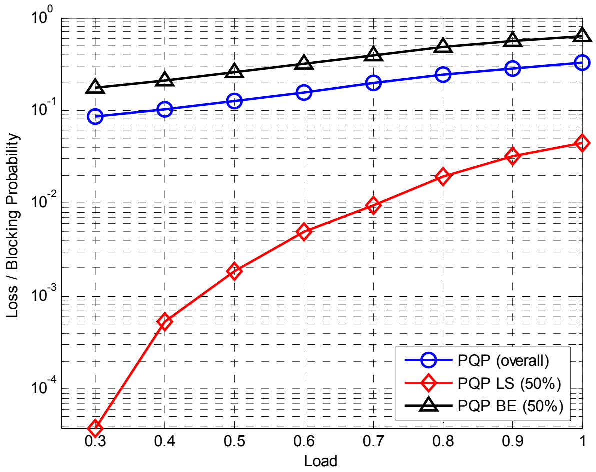


Figure 53. QoS differentiation of synchronous PQP.

Figure 54 shows the blocking probability of LS traffic when varying its proportion of the overall traffic. As expected, whenever the proportion of LS traffic increases its blocking probability increases too. In fact, the mean blocking probability of LS traffic (over all the loads) when its proportion of the overall traffic is 25%, 50%, 75% and 100%, is 1.32x10-4, 1.1x10-2, 5.4x10-2 and 9.7x10-2, respectively. Also, this shows that when proportion of LS traffic is moderate (e.g., 25%), its blocking probability remains under 10-3 which is an interesting performance level for the OBS network under study. It is worth noting that the proportion of LS traffic to be accepted in the OBS network should be carefully determined since increasing the proportion of LS traffic will be at the expense of degrading the performance of BE traffic.

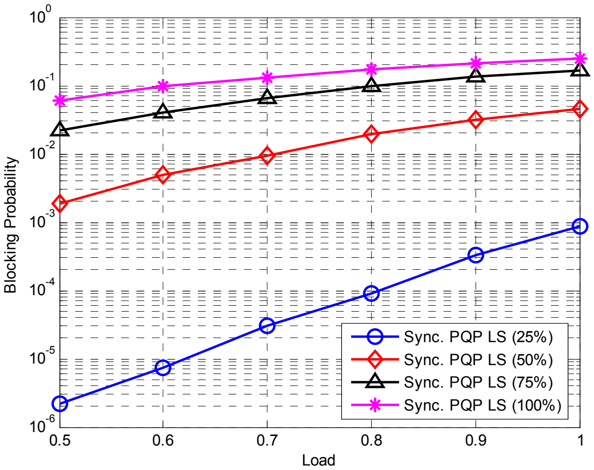
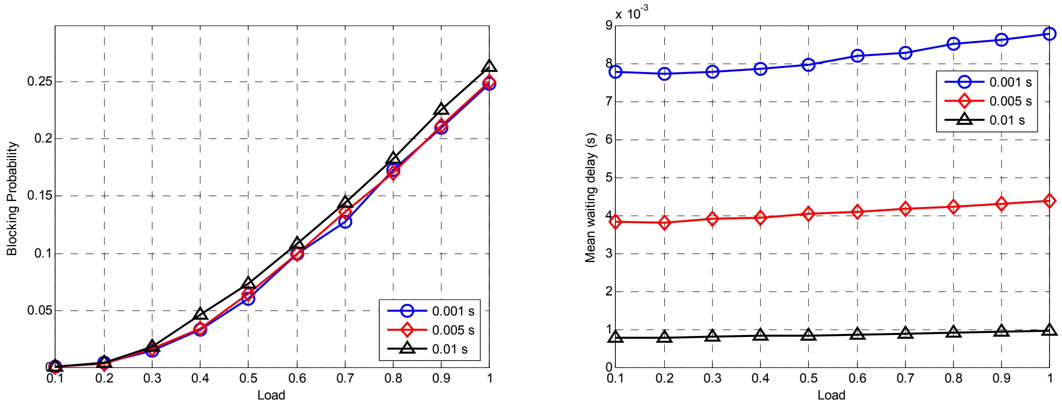


Figure 54. Blocking probability of LS traffic with synchronous PQP.

Figure 55 shows the impact of varying the timeslot duration from 0.01 s to 0.001 s on the blocking probability and the average packet waiting delay. We observe that while reducing the timeslot duration has a negligible impact on the blocking probability; it reduces effectively the mean waiting delay at OBS access nodes. However, determining timeslot duration should be performed while taking into consideration the switching speed of OBS core nodes. Indeed, if the timeslot duration is small compared to the switching time of OBS core nodes, the guard time separating successive timeslots would be bigger than the timeslot duration; this will result in poor resource utilization.



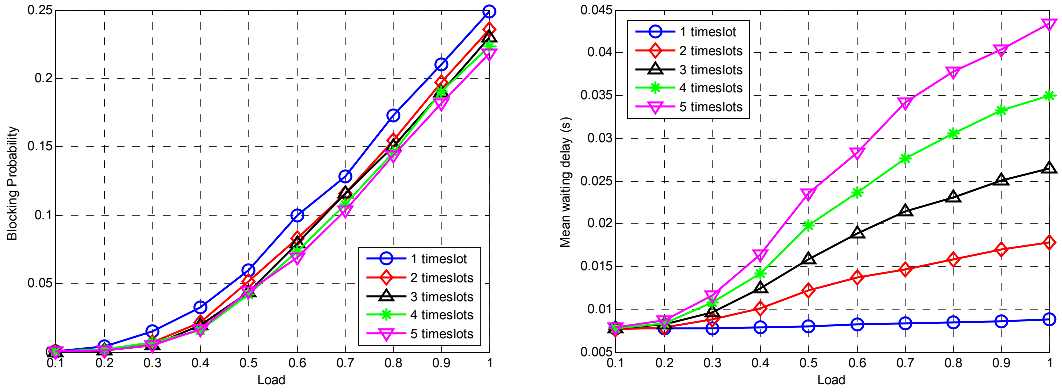
(a) (b)

Figure 55. Impact of varying timeslot duration on loss (a) and delay (b).

Figure 56 shows the impact of increasing buffer size (from 1 timeslot to 5 timeslots) of OBS edge nodes on both blocking probability and delay of LS packets. We observe that increasing buffer size decreases blocking probability at the expense of increasing average packet waiting delay, especially at high loads. In fact, increasing the size of buffers from 1 timeslot to 5 timeslots reduces blocking probability, in average (over all the loads), from 0.097 (for 1 timeslot) to 0.078 (for 5 timeslots); however, this increases average waiting delay from 0.0081 s (for 1 timeslot) to 0.0252 s (for 5 timeslots); these findings show that it is also possible to subdivide guaranteed traffic (LS traffic class) to two subclasses: (a) loss-sensitive class of traffic which tolerates additional delay at the access of the OBS network; and (b) delay sensitive class of traffic (e.g., multi-media traffic) which has to be served with the minimum buffering (i.e., the minimum waiting delay).

5.6.4. Results of PBEWS

As stated in Section 5.3.4, PBEWS can be considered as a general wavelength selection scheme for the classless case (i.e., without QoS provisioning). Hence, we compare the performance of PBEWS to the performance of Best Effort Topology-aware Wavelength Assignment scheme (BETWA) [107] which is a wavelength assignment scheme that outperforms the known wavelength selection schemes for OBS networks, such as Latest Available Unscheduled Channel with Void Filling (LAUC-VF) [32].



(a) (b)

Figure 56. Impact of varying buffer size on loss (a) and delay (b).

Figure 57 shows loss probability of PBEWS and BETWA with 64 wavelengths. We observe that PBEWS outperforms BETWA, especially at low loads. In fact, the mean loss probability of PBEWS (over all the loads) is 0.15 and the mean loss probability of BETWA is 0.27. In addition, whereas loss probability of PBEWS is in the order of 10-4 at traffic load 0.2, it is in the order of 10-2 for BETWA. This shows that PBEWS is able to reduce the loss probability of the OBS network to a remarkable level for the network under study (i.e., without wavelength converters and without fiber delay lines).

Figure 58 shows loss probability of PBEWS and BETWA when fixing the traffic load to 0.5 (corresponding to each number of wavelengths) and varying the number of wavelengths from 16 to 128. We observe that PBEWS is better than BETWA whatever the number of wavelengths. In addition, we can see that both PBEWS and BETWA loss probabilities decrease when the number of wavelengths increases; however, the decrease of PBEWS is more significant than that of BETWA. In fact, with 16 wavelengths, the loss probabilities of PBEWS and BETWA are 0.14 and 0.27, respectively; with 128 wavelengths, the loss probabilities of PBEWS and BETWA are 9.2x10-3 and 0.23, respectively. We conclude that PBEWS is able to exploit the additional bandwidth brought by the increase in the number of wavelengths.

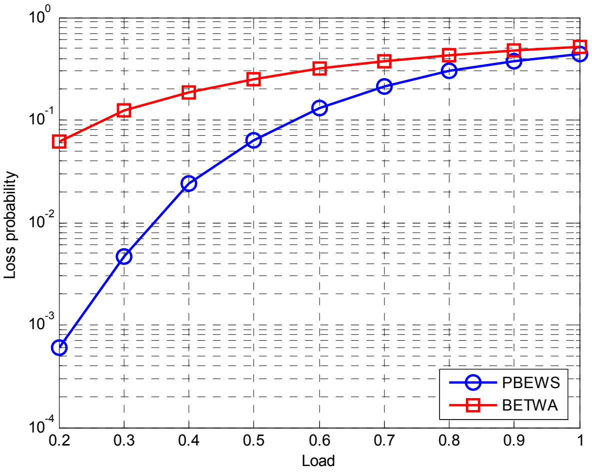


Figure 57. Loss probability of PBEWS vs. BETWA with 64 wavelengths

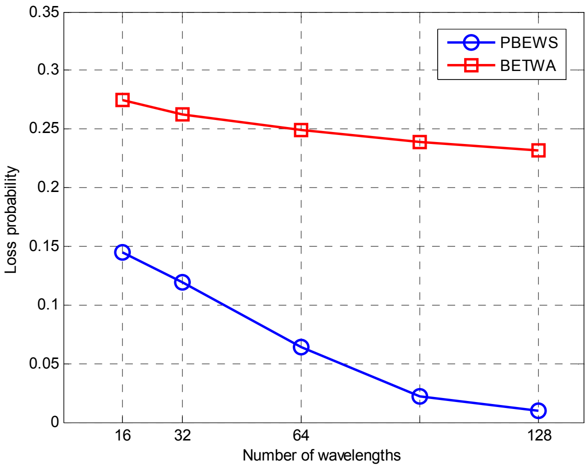


Figure 58. Loss probability of PBEWS vs. BETWA when varying the number of wavelengths.

5.7. Concluding remarks

We proposed a novel QoS provisioning approach (PQP) which offers loss-free transmission inside the OBS network. PQP is based on a routing and wavelength assignment approach (OBSRWA) that reduces the number of required wavelengths to assign a different wavelength to each path in a set of overlapping paths (*asynchronous PQP*); also, PQP is based on a synchronization scheme (PST) to synchronize the transmissions of overlapping paths when the number of available wavelengths in each fiber link does not allow establishing non-overlapping paths (*synchronous PQP*). Furthermore, we proposed a wavelength selection scheme (PBEWS) to improve the performance of best effort traffic and preserve statistical multiplexing and high resource utilization of OBS networks. Numerical results did show that OBSRWA is able to decrease substantially the number of required wavelengths to establish non-overlapping paths; it is also able to reduce the cases of overlaps between paths using a tabu search algorithm when the number of available wavelengths does not allow establishing completely non-overlapping paths. In this case, PST is able not only to synchronize overlapping paths using few and small control packets, but also to maximize the amount of guaranteed traffic that can be carried by each path; this is performed while guaranteeing fairness when allocating bandwidth to conflicting paths. In addition, simulation results using ns-2 simulator did show that asynchronous PQP outperforms AFQD (a QoS provisioning scheme for asynchronous OBS) in terms of loss/blocking probability; also, synchronous PQP outperforms TSA (a QoS provisioning scheme for synchronous OBS) in terms of the number of required wavelengths to guarantee loss-free transmission inside the OBS network. We did show that the wavelength selection scheme PBEWS improves significantly the performance of best effort traffic. These results show clearly that PQP successfully provides absolute QoS provisioning for guaranteed traffic while improving the performance of best effort traffic.

In future work, we plan to propose protocols to make both synchronous PQP and asynchronous PQP adaptive to traffic pattern variations by allowing paths to exchange wavelengths and timeslots.

Chapitre 6 : Conclusion et travaux futurs

Dans ce chapitre, nous passerons en revue les contributions et les résultats de cette thèse et nous dresserons les perspectives pour des travaux de recherche futurs.

6.1. Contributions et résultats de la thèse

Le routage adaptatif et le provisionnement absolu de la qualité de service dans les réseaux OBS ont constitué les deux principaux sujets de cette thèse.

Dans la première contribution (présentée au chapitre 3), nous nous sommes intéressés à l'amélioration des performances des réseaux OBS en termes de la probabilité de perte dans les nœuds du cœur du réseau. Ces pertes sont causées, principalement, par les contentions des longueurs d'onde entre les rafales de données. Ainsi, nous avons exploré la capacité du routage distribué et adaptatif à réduire la probabilité de perte dans le réseau OBS, proactivement, en utilisant le routage multi-chemins et, réactivement, en utilisant le routage alternatif. Ces deux techniques ont été identifiées, depuis plusieurs années, comme étant des techniques efficaces et peu coûteuses pour réduire le nombre de contentions dans le réseau OBS. Notre contribution consistait à proposer une approche basée sur l'apprentissage par renforcement, et plus spécifiquement, l'algorithme du *Q-learning* pour sélectionner les chemins du routage multi-chemins et les liens du routage alternatif (par déflection). Un autre élément clé de cette contribution est le fait de distribuer les décisions du routage sur tous les nœuds du réseau OBS. L'intuition derrière cette démarche est de baser les décisions du routage sur les informations locales de chaque nœud, qui sont plus précises que les informations collectées de bout en bout par les nœuds source. Aussi, cette démarche a permis de réduire le surplus de communication qui est produit par les solutions centralisées ou quasi-centralisées (implantées dans les nœuds du bord) en favorisant la propagation *progressive* de l'information entre les différents nœuds du réseau. Les résultats obtenus confirment que le routage multi-chemins et le routage alternatif sont capables de réduire le niveau de contentions dans le cœur du réseau OBS, en particulier, lorsque la charge du trafic est basse ou modérée. Aussi, l'intégration du routage multi-chemins et du routage alternatif dans une seule approche a amélioré, d'une manière significative, les performances du réseau OBS en termes de la probabilité de perte. En outre, les résultats numériques ont montré que les performances de la solution proposée pour le routage multi-chemins (RLMR) sont meilleures que celles des solutions proposées dans la littérature [36]. Parmi les conséquences de cette réduction de la probabilité de perte, il y a une augmentation du délai de bout en bout des rafales de données, ce qui a été prévu puisque la solution proposée n'utilise pas le routage du plus court chemin qui optimise la métrique du délai. Même si cette augmentation du délai reste modérée et proportionnelle aux délais de propagation des liens du réseau OBS, il sera judicieux de prévoir une solution pour accommoder le trafic sensible aux délais. Un autre problème spécifique aux réseaux OBS, quand la taille des chemins n'est pas déterminée à l'avance, est le temps de décalage insuffisant (*Insufficient offset time*). Même si dans cette thèse nous avons simplifié la solution à ce problème en ajoutant un délai supplémentaire au temps de décalage de base, il sera pertinent de considérer l'intégration des solutions déjà proposées dans la littérature [112, 113] ou même de proposer des solutions spécifiques à l'approche proposée. D'autres problèmes, tels que la livraison en dehors de l'ordre (*Out-of-order delivery*) et la gigue (*Jitter*), sont liés à la nature distribuée de notre solution. En effet, ces problèmes sont dus au fait que les rafales successives peuvent emprunter des chemins différents, et ainsi, arriver à leur nœud destination dans un ordre différent de celui avec lequel elles ont été envoyées dans leur nœud source. Ces problèmes peuvent être amplifiés davantage dans les couches supérieures quand les rafales sont désassemblées et leurs paquets envoyés à leurs réseaux de destination.

La deuxième et la troisième contribution ont été focalisées sur le provisionnement absolu de la qualité de service, et plus spécifiquement, la garantie de la transmission sans pertes des rafales de priorité élevée à l'intérieur du réseau OBS. L'approche générale adoptée dans cette thèse a été celle du groupage des longueurs d'onde (*wavelength* *grouping*) où un ensemble de longueurs d'onde est réservé pour une classe du trafic. Cependant, afin d'atteindre notre objectif de garantir la transmission sans pertes, à l'intérieur du réseau OBS, pour le trafic de priorité élevée, nous avons proposé, dans la deuxième contribution, d'assigner les longueurs d'onde aux nœuds (approche basée sur les nœuds) et nous avons proposé, dans la troisième contribution, d'assigner les longueurs d'onde aux chemins de bout en bout (approche basée sur les chemins). Néanmoins, l'inconvénient de ces approches est le fait de limiter le multiplexage statistique et le niveau d'utilisation élevé des ressources qui caractérisent les réseaux OBS. Pour remédier à ce problème, nous avons utilisé la technique de préemption qui permet aux rafales de type *best effort* d'utiliser toutes les ressources du réseau OBS tout en privilégiant les rafales garanties dans les cas des contentions entre ces derniers et les rafales de type *best effort*. En outre, Pour réduire la probabilité de perte du trafic *best effort*, nous avons proposé deux algorithmes de sélection des longueurs d'onde qui se basent sur l'assignation des longueurs d'onde aux nœuds ou aux chemins. Ces algorithmes peuvent être considérés comme des contributions indépendantes dans le contexte de la sélection des longueurs d'onde dans les réseaux OBS. Les résultats numériques montrent que ces algorithmes réduisent efficacement la probabilité de perte des rafales même dans le cas général (sans provisionnement de qualité de service).

Dans l'approche basée sur les nœuds (présentée au chapitre 4), nous avons proposé un protocole d'échange des longueurs d'onde entre les nœuds pour adapter l'assignation des longueurs d'onde initiale aux variations du modèle du trafic. Ce protocole a permis de réduire le blocage du trafic garanti quand la distribution du trafic sur les différents nœuds du bord n'est pas balancée.

Dans l'approche basée sur les chemins (présentée au chapitre 5), nous avons proposé une approche qui permet de réduire le nombre de longueurs d'onde requis pour établir des chemins sans conflits entre tous les nœuds du bord du réseau OBS. cette approche réduit efficacement le nombre de longueurs d'onde requis en se basant sur le routage des chemins dans le réseaux OBS. Ainsi, le provisionnement absolu de la qualité de service au niveau de chaque chemin dans le réseau devient une solution viable. Par ailleurs, dans le cas où certains chemins se chevauchent à cause du nombre limité de longueurs d'onde dans chaque fibre, nous avons proposé une approche efficace de synchronisation des chemins en conflit qui ne requiert pas d'équipements additionnels.

Les résultats numériques ont montré que, dans certains scénarios, l'approche basée sur les chemins est meilleure que l'approche basée sur les nœuds en termes de la quantité du trafic garanti bloquée à l'accès du réseau ainsi qu'en termes de la probabilité de perte du trafic *best effort*. Néanmoins, cette comparaison n'est pas toujours directe puisqu'elle dépend du nombre de nœuds dans le réseau OBS et du nombre requis de longueurs d'onde pour établir des chemins sans conflits. Ainsi, si par exemple, le nombre de longueurs d'onde est inférieur au nombre requis pour établir des chemins sans conflits et que le fonctionnement synchrone de certains chemins n'est pas souhaité, l'approche basée sur les nœuds peut être la solution la plus appropriée.

6.2. Perspectives et travaux futurs

Les contributions proposées dans cette thèse ouvrent plusieurs pistes de recherche pour des travaux futurs.

Nous avons déjà souligné, dans la section précédente, les problèmes liés au routage adaptatif. Parmi ces problèmes, il y a le temps de décalage insuffisant qui est spécifique aux réseaux OBS et les problèmes liés au routage distribué tels que la livraison en dehors de l'ordre (*Out-of-order delivery*) et la gigue (*jitter*). Ainsi, trouver des solutions à ses problèmes dans le contexte de l'approche du routage adaptatif proposée dans le chapitre 3 rendra cette approche plus robuste.

Dans les contributions sur l'approvisionnement absolu de la qualité de service, nous avons supposé l'existence de deux classes de service : une classe garantie (sensible aux pertes) et une classe *best effort*. Il sera, donc, intéressant de considérer le cas plus général où plusieurs classes de trafic coexistent. En particulier, dans l'approche basée sur les chemins, nous avons identifié un cas prometteur où les classes de trafic sensibles aux délais peuvent être accommodées avec le minimum de délai possible tout en continuant de leur garantir le provisionnement absolu de la qualité de service en termes de la probabilité de perte. Par ailleurs, dans l'approche basée sur les chemins, nous n'avons pas proposé une solution pour prendre en considération le trafic non balancé dans l'assignation des longueurs d'onde et des intervalles de temps (*timeslots*) aux chemins. Un protocole similaire à celui proposé pour l'approche basée sur les nœuds et qui opère aussi au niveau des intervalles de temps rendra cette approche plus robuste et adaptative aux variations du modèle du trafic. En outre, que ce soit dans l'approche basée sur les nœuds ou dans l'approche basée sur les chemins, nous avons adopté la sélection des longueurs d'onde comme approche proactive pour réduire la probabilité de perte du trafic *best effort*. Ainsi, l'utilisation des mécanismes réactifs, tel que le routage alternatif ou la conversion des longueurs d'onde clairsemée, conjointement avec la sélection des longueurs d'onde pourra améliorer de manière significative les performances de cette classe de trafic.

Enfin, considérer le réseau OBS et les solutions proposées dans cette thèse dans le contexte plus général des réseaux de nouvelle génération, constituera, sans doute, une piste de recherche prometteuse. En effet, nous avons déjà identifié la convergence des réseaux optiques et des réseaux sans fil, et plus spécifiquement, la convergence des réseaux OBS et des réseaux maillés sans fil (*Wireless Mesh Networks*) [114] comme étant un cadre propice à l'extension des travaux de recherche présentés dans cette thèse.

Bibliographie

[1] J. Zheng and H. T. Mouftah, *Optical WDM Networks : Concepts and Design Principles*, New York, IEEE Press/Wiley and Sons. 2004.

[2] M. Maier, *Optical Switching Networks*, Cambridge, UK, Cambridge University Press. 2008.

[3] T. Battestilli and H. Perros, *An introduction to optical burst switching.* IEEE Communications Magazine, Vol. 41(8): p. S10 - S15, 2003.

[4] I. Chlamtac, A. Ganz, and G. Karmi, *Lightpath communications: an approach to high bandwidth optical WAN's.* IEEE Transactions on Communications, Vol. 40(7): p. 1171 - 1182, 1992.

[5] M. J. O'Mahony, D. Simeonidou, D. K. Hunter, and A. Tzanakaki, *The application of optical packet switching in future communication networks.* IEEE Communications Magazine, Vol. 39(3): p. 128 - 135, 2001.

[6] J. M. Finochietto, J. Aracil, A. Ferreiro, J. P. Fernandez-Palacios Gimenez, and O. Gonzalez de Dios, *Migration Strategies Toward All Optical Metropolitan Access Rings.* IEEE/OSA Journal of Lightwave Technology, Vol. 25(8): p. 1918-1930, 2007.

[7] A. Belbekkouche, J. Rezgui, and A. Hafid. *QoS Provisioning for Wireless Mesh and Optical Burst Switching Convergence*. in Proceedings of *IEEE Wireless Communications & Networking Conference (WCNC'10)*, Sydney, Australia, 2010.

[8] A. Belbekkouche, A. Hafid, and M. Gendreau, *Novel reinforcement learning-based approaches to reduce loss probability in buffer-less OBS networks.* Computer Networks, Vol. 53(12): p. 2091-2105, 2009.

[9] A. Belbekkouche, A. Hafid, M. Tagmouti, and M. Gendreau, *Topology-aware wavelength partitioning for DWDM OBS networks: A novel approach for absolute QoS provisioning.* Computer Networks, Vol. 54(18): p. 3264-3279, 2010.

[10] J. Ramamirtham and J. Turner. *Time sliced optical burst switching*. in Proceedings of *Twenty-second Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'03)*, San Francisco, CA, USA, 2003.

[11] S. Sheeshia and C. Qiao. *Synchronous Optical Burst Switching*. in Proceedings of *First International Conference on Broadband Networks (BroadNets'04)*, San José, CA, USA, 2004.

[12] A. Rugsachart and R. A. Thompson. *An analysis of time-synchronized optical burst switching*. *Workshop on High Performance Switching and Routing (HPSR'06)*, Poznañ, Poland, 2006.

[13] O. Ozturk, E. Karasan, and N. Akar, *Performance Evaluation of Slotted Optical Burst Switching Systems With Quality of Service Differentiation.* IEEE/OSA Journal of Lightwave Technology, Vol. 27(14): p. 2621 - 2633, 2009.

[14] H. Jeong, J. Choi, and M. Kang, *A tree-based slot allocation algorithm for loss-free slotted OBS networks.* Photonic Network Communications, Vol. 19(3): p. 320-330, 2010.

[15] Z. Zhang, L. Liu, and Y. Yang, *Slotted Optical Burst Switching (SOBS) networks.* Computer Communications, Vol. 30(18): p. 3471-3479, 2007.

[16] T.-W. Um, J. K. Choi, J. Guo, R. Won, and B. S. Lee, *Soft-State Bandwidth Reservation Mechanism for Slotted Optical Burst Switching Networks.* ETRI Journal, Vol. 30(2): p. 216-226, 2008.

[17] A. Belbekkouche, A. Hafid, M. Gendreau, and M. Tagmouti, *Path-based QoS Provisioning for Optical Burst Switching Networks.* Submitted to IEEE/ACM Transactions on Networking, Vol.,

[18] C. Qiao and M. Yoo, *Optical burst switching (OBS) - a new paradigm for an optical Internet.* Journal of High Speed Networks, Vol. 8(1): p. 69-84, 1999.

[19] C. Qiao and M. Yoo, *Choices, features and issues in optical burst switching.* Optical Networks Magazine, Vol. 1(2): p. 36-44, 2000.

[20] C. Gauger. *Novel network architecture for optical burst transport*. *Communication networks and computer engineering report N° 92-2006*, University of Stuttgart, 2006.

[21] T. Battestilli and H. Perros. *Optical Burst Switching: a Survey*. *Technical Report, UMI Order Number TR-2002-10*, North Carolina State University at Raleigh, 2002.

[22] V. Eramo, A. Germoni, C. Raffaelli, and M. Savi, *Packet loss analysis of shared-per-wavelength next term multi-fiber all-optical switch with parallel scheduling.* Computer Networks, Vol. 53(2): p. 202-216, 2009.

[23] J. Choi, J. Choi, and M. Kang, *Dimensioning burst assembly process in optical burst switching networks.* IEICE Transactions on Communications, Vol. E88-B(10): p. 3855–3863, 2005.

[24] X. Yu, Y. Chen, and C. Qiao. *A study of traffic statistics of assembled burst traffic in optical burst switched networks*. in Proceedings of *Optical Networking and Communications Conference (OptiComm)*, Boston, USA, 2002.

[25] K. Laevens. *Traffic characteristics inside optical burst switched networks*. in Proceedings of *Optical Networking and Communications Conference (OptiComm)*, Boston, USA, 2002.

[26] V. M. Vokkarane, Q. Zhang, J. P. Jue, and B. Chen. *Generalized burst assembly and scheduling techniques for QoS support in optical burst-switched networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'02)*, Taipei, Taiwan, 2002.

[27] A. Ge, F. Callegati, and L. S. Tamil, *On optical burst switching and self-similar traffic.* IEEE Communications Letters, Vol. 4(3), 2000.

[28] S. Verma, H. Chaskar, and R. Ravikanth, *Optical burst switching: a viable solution for terabit IP backbone.* IEEE Network, Vol. 14(6): p. 48-53, 2000.

[29] S.-Y. Oh, H. H. Hong, and M. Kang, *A data burst assembly algorithm in optical burst switching networks.* ETRI Journal, Vol. 24(4): p. 311-322, 2002.

[30] X. Cao, J. Li, Y. Chen, and C. Qiao. *Assembling TCP/IP packets in optical burst switched networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'02)*, Taipei, Taiwan, 2002.

[31] J. Y. Wei, J. L. Pastor, R. S. Ramamurthy, and Y. Tsai, *Just-in-time optical burst switching for multi-wavelength networks*, in *Broadband communications: convergence of network technologies*1999, Kluwer Academic Publishers: Norwell, MA, USA. p. 339-352.

[32] Y. Xiong, M. Vandenhoute, and H. C. Cankaya, *Control architecture in optical burst-switched WDM networks.* IEEE Journal on Selected Areas in Communications, Vol. 18(10): p. 1838-1851, 2000.

[33] J. Li, C. Qiao, and Y. Chen, *Recent progress in the scheduling algorithms in optical-burst-switched networks.* Journal of Optical Networking, Vol. 3: p. 229-241, 2004.

[34] S. Yao, B. Mukherjee, S. J. B. Yoo, and S. Dixit. *All-optical packet-switched networks: a study of contention-resolution schemes in an irregular mesh network with variable-sized packets*. in Proceedings of *Optical Networking and Communications Conference (OptiComm)*, Richardson, TX, USA, 2000.

[35] S. Dixit, *IP over WDM: Building the Next-Generation Optical Internet*, Wiley. 2003.

[36] Y. V. Kiran, T. Venkatesh, and C. S. R. Murthy, *A Reinforcement Learning Framework for Path Selection and Wavelength Selection in Optical Burst Switched Networks.* IEEE Journal on Selected Areas in Communications, Vol. 25(9): p. 18-26, 2007.

[37] G. R. V. Thodime, V. M. Vokkarane, and J. P. Jue. *Dynamic congestion-based load balanced routing in optical burst-switched networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'03)*, San Francisco, USA, 2003.

[38] L. Yang and G. N. Rouskas, *Adaptive path selection in OBS networks.* IEEE/OSA Journal of Lightwave Technology, Vol. 24(8): p. 3002 - 3011, 2006.

[39] D. Ishii, N. Yamanaka, and I. Sasase, *Self-learning Route Selection Scheme using Multipath Searching Packets in OBS Network.* Journal of Optical Networking, Vol. 4(7): p. 432-445, 2005.

[40] S. Ganguly, S. Bhatnagar, R. Izmailov, and Chunming Qiao. *Multi-path adaptive optical burst forwarding*. *Workshop on High Performance Switching and Routing (HPSR'04)*, Phoenix, Arizona, USA, 2004.

[41] J. Lu, Y. Liu, M. Gurusamy, and K. C. Chua. *Gradient projection based multi-path traffic routing in optical burst switching networks*. *Workshop on High Performance Switching and Routing (HPSR'06)*, Poznañ, Poland, 2006.

[42] J. Teng and G. N. Rouskas, *A Traffic Engineering Approach to Path Selection in OBS Networks.* OSA Journal of Optical Networking, Vol. 4(11): p. 759-777, 2005.

[43] M. Klinkowski, M. Pioro, D. Careglio, M. Marciniak, and J. Sole-Pareta, *Non-linear optimization for multi-path source routing in obs networks.* IEEE Communications Letters, Vol. 11(12): p. 1016 - 1018, 2007.

[44] J. Li, G. Mohan, and K. C. Chua, *Dynamic load balancing in ip-over-wdm optical burst switching networks.* Computer Networks, Vol. 47(3): p. 393-408, 2005.

[45] Y. Hirota, H. Tode, and K. Murakami. *Cooperation method considering wavelength assignment and routing problem in optical burst switched networks*. *Optical Fiber Communication Conference and Exhibit (OFC'06)*, Anaheim, CA, USA, 2006.

[46] C. G. Argos, O.G. de Dios, and J. Aracil. *Adaptive Multi-Path Routing for OBS Networks*. in Proceedings of *9th International Conference on Transparent Optical Networks (ICTON'07)*, Rome, Italy, 2007.

[47] 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'02)*, New York, USA, 2002.

[48] X. Wang, H. Morikawa, and T. Ayoama, *Burst optical deflection routing protocol for wavelength routing WDM networks.* Optical Networks Magazine, Vol. 3(6): p. 12-19, 2002.

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

[50] Y. Chen, H. Wu, D. Hu, and C. Qiao. *Performance analysis of optical burst switched node with deflection routing*. in Proceedings of *IEEE International Conference on Communications (ICC'03)*, Anchorage, USA, 2003.

[51] 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 Conference of the IEEE Communications Society (INFOCOM'04)*, Hong Kong, China, 2004.

[52] C. Cameron, A. Zalesky, and M. Zukerman, *Prioritized Deflection Routing in Optical Burst Switching Networks.* IEICE Transactions on Communication, Vol. E88-B(5): p. 1861-1867, 2005.

[53] S. Lee, K. Sriram, H. Kim, and J. Song, *Contention-Based Limited Deflection Routing Protocol in Optical Burst-Switched Networks.* IEEE Journal on Selected Areas in Communications, Vol. 23(8): p. 1596- 1611, 2005.

[54] S. Lee, L. Kim, J. Song, D. Griffith, and K. Sriram, *Dynamic deflection routing with virtual wavelength assignment in optical burst-switched networks.* Photonic Network Communications, Vol. 9(3): p. 347-356, 2005.

[55] T. Coutelen, H. Elbiaze, B. Jaumard, and A. Metnani. *Measurement-based alternative routing strategies in optical burst-switched networks*. in Proceedings of *7th International Conference on Transparent Optical Networks (ICTON'05)*, Barcelona, Spain 2005.

[56] T. Coutelen, H. Elbiaze, B. Jaumard, and A. Metnani. *Impact of Network Topology and Traffic Load Balancing on Deflection Routing Efficiency*. in Proceedings of *IASTED International Conference on Optical Communication Systems and Networks (OCSN'05)*, Banff, Canada, 2005.

[57] N. Ogino and H. Tanaka, *Deflection Routing for Optical Bursts Considering Possibility of Contention at Downstream Nodes.* IEICE Transactions on Communication, Vol. E88-B(9): p. 3660-3667, 2005.

[58] Y. Du, C. Zhu, X. Zheng, Y. Guo, and H. Zhang. *A Novel Load Balancing Deflection Routing Strategy in Optical Burst Switching Networks*. *Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference (OFC/NFOEC'07)*, Anaheim, CA, USA, 2007.

[59] V. M. Vokkarane and J. P. Jue. *Prioritized routing and burst segmentation for QoS in optical burst-switched networks*. *Optical Fiber Communication Conference and Exhibit (OFC'02)*, Anaheim, CA, USA, 2002.

[60] Y. Zhu, G. N. Rouskas, and H. G. Perros, *A Comparison of Allocation Policies in Wavelength Routing Networks.* Photonic Networks Communications, Vol. 2(3): p. 265-293, 2000.

[61] J. Teng and G. N. Rouskas. *On Wavelength Assignment in Optical Burst Switched Networks*. in Proceedings of *First International Conference on Broadband Networks (BroadNets'04)*, San José, CA, USA, 2004.

[62] X. Wang, H. Morikawa, and T. Aoyama, *Priority-based wavelength assignment algorithm for burst switched WDM optical networks.* IEICE Transactions on Communications, Vol. E86-B(5): p. 1508-1514, 2003.

[63] M. Yoo, C. Qiao, and S. Dixit, *QoS performance of optical burst switching in IP-over-WDM networks.* IEEE Journal on Selected Areas in Communications, Vol. 10(10): p. 2062-2071, 2000.

[64] 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, Vol. 21(7): p. 1198 - 1209, 2003.

[65] V. M. Vokkarane and J. P. Jue, *Segmentation-based nonpreemptive channel scheduling algorithms for optical burst-switched networks.* IEEE/OSA Journal of Lightwave Technology, Vol. 23(10): p. 3125 - 3137, 2005.

[66] L. Yang, Y. Jiang, and S. Jiang. *A probabilistic preemptive scheme for providing service differentiation in OBS networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'03)*, San Francisco, USA, 2003.

[67] C. W. Tan, G. Mohan, and J. C.-S. Lui, *Achieving Multi-Class Service Differentiation in WDM Optical Burst Switching Networks: A Probabilistic Preemptive Burst Segmentation Scheme.* IEEE Journal on Selected Areas in Communications, Vol. 24(12): p. 106 - 119, 2006.

[68] Y. Wang and B. Ramamurthy. *CPQ: A Control Packet Queuing Optical Burst Switching Protocol for Supporting QoS*. *The Third International Workshop on Optical Burst Switching*, San José, CA, USA, 2004.

[69] F. Farahmand and J. P. Jue. *Supporting QoS with look-ahead window contention resolution in optical burst switched networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'03)*, San Francisco, USA, 2003.

[70] Y. Chen, M. Hamdi, and D. H. K. Tsang. *Proportional QoS over OBS networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'01)*, San Antonio, TX, USA, 2001.

[71] J. Phuritatkul and Y. Ji, *Resource Allocation Algorithms for Controllable Service Differentiation.* IEICE Transactions on Communications, Vol. E88-B(4): p. 1424–1431, 2005.

[72] B. Zhou, M. Bassiouni, and G. Li, *Routing and Wavelength Assignment in Optical Networks Using Logical Link Representation and Efficient Bitwise Computation.* Journal of Photonic Network Communications, Vol. 10(3): p. 333–346, 2005.

[73] W. Liao and C.-H. Loi, *Providing service differentiation for optical-burst-switched networks.* IEEE/OSA Journal of Lightwave Technology, Vol. 22(7): p. 1651–1660, 2004.

[74] 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, Vol. 22(9): p. 1781-1795, 2004.

[75] S. Kim, J. S. Choi, and M. Kang, *Providing absolute differentiated services for optical burst switching networks: loss differentiation.* IEE Proceedings-Communications, Vol. 152(4): p. 439-446, 2005.

[76] J. Phuritatkul, Y. Ji, and S. Yamada, *Proactive Wavelength Pre-Emption for Supporting Absolute QoS in Optical-Burst-Switched Networks.* IEEE/OSA Journal of Lightwave Technology, Vol. 25(5): p. 1130 -1137, 2007.

[77] X. Guan, I. L.-J. Thng, Y. Jiang, and H. Li, *Providing absolute QoS through virtual channel reservation in optical burst switching networks.* Computer Communications, Vol. 28(9): p. 967-986, 2005.

[78] M. H. Phùng, K. C. Chua, G. Mohan, M. Motani, and T. C. Wong, *An Absolute QoS Framework for Loss Guarantees in Optical Burst-Switched Networks.* IEEE Transactions on Communications, Vol. 55(6): p. 1191 - 1201, 2007.

[79] L. Hongbo and H. T. Mouftah. *A New Absolute QoS Differentiation Scheme Supporting Best-Effort Class in OBS Networks*. in Proceedings of *9th International Conference on Transparent Optical Networks (ICTON'07)*, Rome, Italy, 2007.

[80] L. Hongbo and H. T. Mouftah. *Absolute QoS Differentiation with Best-Effort Class Support in Optical Burst Switching Networks*. in Proceedings of *IEEE Symposium on Computers and Communications (ISCC'07)*, Aveiro, Portugal, 2007.

[81] L. Yang and G. N. Rouskas. *A Framework for Absolute QoS Guarantees in Optical Burst Switched Networks*. in Proceedings of *Third International Conference on Broadband Communications, Networks, and Systems (Broadnets'06)*, San José, CA, USA, 2006.

[82] K. C. Chua, M. Gurusamy, Y. Liu, and M. H. Phung, *Quality of Service in Optical Burst Switched Networks*, Springer. 2007.

[83] N. Akar, Ezhan Karasan, K. G. Vlachos, E. A. Varvarigos, D. Careglio, M. Klinkowski, and J. Solé-Pareta, *A survey of quality of service differentiation mechanisms for optical burst switching networks.* Optical Switching and Networking, Vol. 7(1): p. 1-11, 2010.

[84] L. Kleinrock, *Queueing Systems, Volume 1: Theory*, New York, USA, Wiley Interscience. 1975.

[85] L. Yang and G. N. Rouskas. *Path Switching in Optical Burst Switched Networks*. in Proceedings of *Networking*, Waterloo, Canada, 2005.

[86] A. Zalesky, Hai Le Vu, M. Zukerman, Z. Rosberg, and E.W.M. Wong. *Evaluation of limited wavelength conversion and deflection routing as methods to reduce blocking probability in optical burst switched networks*. in Proceedings of *IEEE International Conference on Communications (ICC'04)*, Paris, France, 2004.

[87] A. Belbekkouche, A. Hafid, and M. Gendreau. *A Reinforcement Learning-based Deflection Routing Scheme for buffer-less OBS Networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'08)*, New Orleans, USA, 2008.

[88] C.-F. Hsu and T.-L. Liu, *On deflection routing in optical burst-switched networks.* Journal of High Speed Networks, Vol. 14(4): p. 341-362, 2005.

[89] K. Littman and M. Moore, *Reinforcement Learning: A survey.* Journal of Artificial Intelligence Research, Vol. 4: p. 237-285, 1996.

[90] P. D. Christopher and J. C. H. Watkins, *Q-Learning.* Machine Learning, Vol. 8: p. 279-292, 1992.

[91] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, Cambridge, Massachusetts, The MIT Press. 1998.

[92] J. P. Jue. *An algorithm for loopless deflection in photonic packet-switched networks*. in Proceedings of *IEEE International Conference on Communications (ICC'02)*, New York, USA, 2002.

[93] *"NS-Simulator", www.isi.edu/nsnam/ns*.

[94] S. Gowda, R. K. Shenai, K. M. Sivalingam, and H. C. Cankaya. *Performance evaluation of TCP over optical burst-switched (OBS) WDM networks*. in Proceedings of *IEEE International Conference on Communications (ICC'03)*, Anchorage, USA, 2003.

[95] H. Guo, Z. Lan, J. Wu, Z. Gao, X. Li, J. Lin, and Y. Ji. *A testbed for optical burst switching network*. in Proceedings of *Optical Fiber Communication Conference and Exposition and The National Fiber Optic Engineers Conference (OFC/NFOEC'05)*, Anaheim, CA, USA, 2005.

[96] Y. Sun, T. Hashiguchi, V. Q. Minh, X. Wang, H. Morikawa, and T. Aoyama, *Design and implementation of an optical burst-switched network testbed.* IEEE Communications Magazine, Vol. 43(11): p. S48-S55, 2005.

[97] *"EtherBurst optical switch", www.matissenetworks.com*.

[98] *"IBM ILOG CPLEX Optimizer 10.11", www-01.ibm.com/software/integration/optimization/cplex-optimizer/*.

[99] P. M. Pardalos, F. Rendl, and H. Wolkowicz. *The Quadratic Assignment Problem: A Survey and Recent Developments. the DIMACS Workshop on Quadratic Assignment Problems, volume 16 of DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 1994.

[100] S. Sahni and T. Gonzalez, *P-complete approximation problems.* Journal of the Association of Computing Machinery, Vol. 23: p. 555-565, 1976.

[101] É. D. Taillard, *Robust taboo search for the quadratic assignment problem.* Parallel Computing, Vol. 17: p. 443-455, 1991.

[102] F. Glover, *Future paths for Integer Programming and Links to Artificial Intelligence.* Computers and Operations Research, Vol. 5: p. 533-549, 1986.

[103] J. Teng and G. N. Rouskas. *Routing path optimization in optical burst switched networks*. in Proceedings of *9th Conference on Optical Network Design and Modelling (ONDM'05)*, Milan, Italy, 2005.

[104] J. Berthold, A. A .M. Saleh, L. Blair, and J. M. Simmons, *Optical Networking: Past, Present, and Future.* IEEE/OSA Journal of Lightwave Technology, Vol. 26(9): p. 1104 - 1118, 2008.

[105] *"Next Generation Network Global Standards Initiative (NGN-GSI)", http://www.itu.int/ITU-T/ngn/.*

[106] S. Orlowski, M. Pióro, A. Tomaszewski, and R. Wessäly. *SNDlib 1.0-Survivable Network Design Library*. in Proceedings of *The 3rd International Network Optimization Conference (INOC)*, Spa, Belgium, 2007.

[107] A. Belbekkouche, A. Hafid, M. Tagmouti, and M. Gendreau. *An Absolute and Fair QoS Differentiation Scheme For DWDM OBS Networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'09)*, Honolulu, USA, 2009.

[108] M. Garey and D. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Freeman & Co. 1979.

[109] Z. Wang and J. Crowcroft, *Quality-of-service routing for supporting multimedia applications.* IEEE Journal on Selected Areas in Communications, Vol. 14(7): p. 1228 - 1234, 1996.

[110] E. Malaguti and P. Toth, *A survey on vertex coloring problems.* International Transactions in Operational Research, Vol. 17(1): p. 1-34, 2009.

[111] D. Brélaz, *New methods to color the vertices of a graph.* Communications of the ACM Vol. 22(4): p. 251-256, 1979.

[112] A. Belbekkouche and A. Hafid. *An Adaptive Reinforcement Learning-based Approach to Reduce Blocking Probability in Buffer-less OBS Networks*. in Proceedings of *IEEE International Conference on Communications (ICC'07)*, Glasgow, UK, 2007.

[113] T. Coutelen, H. Elbiaze, and B. Jaumard. *An efficient adaptive offset mechanism to reduce burst losses in OBS networks*. in Proceedings of *IEEE Global Telecommunications Conference (GLOBECOM'05)*, St. Louis, USA, 2005.

[114] A. Belbekkouche, J. Rezgui, and A. Hafid, *Wireless mesh and optical burst switching convergence for a novel metropolitan area network architecture.* To appear in Computer Networks (accepted August 2010), doi:10.1016/j.comnet.2010.08.006.