

Université de Montréal

Accès et routage optique en mode de commutation de rafales

par

Thomas Coutelen

Département d'informatique et
de recherche opérationnelle

Faculté des arts et sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M. Sc.)
en informatique

Août 2005

Copyright © Thomas Coutelen, 2005



QA

76

U54

2005

V. 049

AVIS

L'auteur a autorisé l'Université de Montréal à reproduire et diffuser, en totalité ou en partie, par quelque moyen que ce soit et sur quelque support que ce soit, et exclusivement à des fins non lucratives d'enseignement et de recherche, des copies de ce mémoire ou de cette thèse.

L'auteur et les coauteurs le cas échéant conservent la propriété du droit d'auteur et des droits moraux qui protègent ce document. Ni la thèse ou le mémoire, ni des extraits substantiels de ce document, ne doivent être imprimés ou autrement reproduits sans l'autorisation de l'auteur.

Afin de se conformer à la Loi canadienne sur la protection des renseignements personnels, quelques formulaires secondaires, coordonnées ou signatures intégrées au texte ont pu être enlevés de ce document. Bien que cela ait pu affecter la pagination, il n'y a aucun contenu manquant.

NOTICE

The author of this thesis or dissertation has granted a nonexclusive license allowing Université de Montréal to reproduce and publish the document, in part or in whole, and in any format, solely for noncommercial educational and research purposes.

The author and co-authors if applicable retain copyright ownership and moral rights in this document. Neither the whole thesis or dissertation, nor substantial extracts from it, may be printed or otherwise reproduced without the author's permission.

In compliance with the Canadian Privacy Act some supporting forms, contact information or signatures may have been removed from the document. While this may affect the document page count, it does not represent any loss of content from the document.

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

Accès et routage optique en mode de commutation de rafales

présenté par :

Thomas Coutelen

a été évalué par un jury composé des personnes suivantes :

Abdelhakim Hafid
(président-rapporteur)

Brigitte Jaumard
(directeur de maîtrise)

Halima Elbiaze
(codirecteur de maîtrise)

Michel Gendreau
(examineur)

Mémoire accepté le 31 octobre 2005

Sommaire

Le médium optique est un acteur important en télécommunications de par la bande passante qu'il met en jeu. L'absence de mémoire optique est cependant la contrainte majeure pour la mise en place d'une telle technologie. Il est donc nécessaire de concevoir des protocoles de routage adaptés pour utiliser au mieux la capacité offerte par le médium tout en ce pliant à cette contrainte matérielle.

La commutation de circuits prône pour le moment, car elle respecte les contraintes matérielles ; elle souffre cependant d'une mauvaise utilisation des ressources. Une alternative orientée paquets est aujourd'hui encore difficile à concevoir pour des raisons technologiques. La commutation par rafales quant à elle semble optimiser de façon non négligeable l'utilisation de la bande passante et ne nécessite pas de mémoire optique.

Elle consiste à envoyer les données sous forme de gros paquets ou rafales (*burst*) précédés par un entête de signalisation. Le temps (*offset time*, OT) séparant l'entête des données sert à la conversion et au traitement de l'entête à chaque noeud intermédiaire. Dès lors, la mémoire optique n'est plus nécessaire pour le routage des données. Le principal défi réside alors dans la difficulté de la résolution des contentions qui peuvent engendrer un fort taux de perte.

Dans ce rapport, nous nous proposons d'étudier l'accès au réseau et le routage dans le réseau. Pour l'accès, nous comparons plusieurs métriques pour choisir la longueur d'onde de la rafale et proposons une méthode adaptative pour le calcul de l'OT. La méthode adaptative permet d'améliorer le compromis entre le taux de perte et le délai et surtout résoud le problème d'OT pour des trafics variables. Pour le routage, nous étudions le routage par déflexion consistant à dévier une rafale en contention sur un chemin alternatif. Une approche orientée équilibrage de charge est proposée pour l'initialisation des tables de routage.

Nous testons les contributions de ce projet par simulation sur le réseau optique européen EONET du projet COST266 et les validons en observant le taux de perte et le délai des trafics mis en jeu. Les résultats montrent la pertinence du modèle de répartition des charges proposé par une diminution des pertes et des délais. L'algorithme adaptatif quant à lui permet de réduire les délais à faible charge et le taux de perte à forte charge.

Mots Clés : *fibres optiques, WDM, commutation OBS, offset time, contention, qualité de service*

Abstract

Optical medium is an important part of the new telecommunication orientation because of the bandwidth it offers. The lack of optical memory is the main constraint we have to deal with to build an optical network. Thus, we have to design new routing protocols to fit the constraints and use their potential in the best way.

Today, Optical Circuit Switching is the only used paradigm because it does not need optical memory but it suffers from a clear bandwidth underutilization. Another philosophy stands in an Optical Packet alternative but it is still hard to design because of the immaturity of the hardware.

Optical Burst Switching seems to optimize resource utilization with no need of special hardware (especially no optical memory). It consists in sending data in a big packet (burst) with a signalization header that leaves ahead of it. The Offset Time between the header and the burst is used to process the header and reserve the bandwidth so that the switches are preset before the incoming of the burst : the data can be switched without conversion or memory. The main issue takes place in the contention resolution that can lead to a huge loss rate.

In this study, we worked on the network access and the routing inside the network. For the access point of view, we propose several metrics to choose the wavelength and an adaptive definition of the offset time. On the routing side of the problem, we used deflection as the only feature to resolve contention. It consists in sending one of the two contending bursts on an alternate available wavelength. We finally present a model to initialize the routing tables with load balancing consideration.

We tested our proposals through simulation on the European network and the traffic instance of the cost_266 project. The quality and comparison metrics are the loss rate and the delay of the incoming traffics.

We validate our proposals with the use of the load balancing model which improves both the delay and the lossrate. The adaptative offset time definition reduces the delay under low network load and improve the lossrate under high network load.

Keywords : *Optical Fibers, WDM, OBS, Offset Time, Contention, Quality of Service*

Table des matières

Sommaire	iv
Abstract	vi
Table des matières	vii
Liste des tableaux	xii
Table des figures	xiii
Liste des algorithmes	xv
Liste des sigles et abréviations	xvii
Remerciements	xviii
Dédicaces	xix
1 Introduction	1
1.1 Problématique	3
1.2 Contributions	4
2 Architecture d'un réseau optique	5
2.1 Motivations pour l'utilisation de la lumière comme support d'information	5

2.2	Les fibres optiques	6
2.2.1	La fibre multimode	7
2.2.2	La fibre monomode	7
2.3	Les commutateurs optiques	7
2.3.1	Les MEMS (<i>MicroElectroMechanical Systems</i>)	8
2.3.2	Les SOA (<i>Semiconductor Optical Amplifier</i>)	9
2.4	WDM	10
2.5	Caractéristiques actuelles des transmissions optiques	11
2.6	Réseau IPoWDM	11
2.7	La commutation dans un réseau optique	12
2.7.1	Introduction	12
2.7.2	Commutation de paquets optiques (OPS)	12
2.7.3	Commutation de circuits optiques (OCS)	13
2.7.4	Commutation de rafales (OBS)	15
2.7.5	Conclusions	17
3	Architecture des réseaux OBS	19
3.1	Le réseau	19
3.2	Les routeurs OBS	19
3.2.1	Module électro-optique	20
3.2.2	Module tout optique	22

3.2.3	Informations pour le routage	23
3.3	Création des rafales	23
3.3.1	Architecture du module de création de rafales	24
3.3.2	Déclenchement d'une rafale	24
3.4	Signalisation dans un réseau OBS	25
3.4.1	Classification des protocoles de signalisation	25
3.4.2	Protocoles de signalisation OBS sans acquittement (JIT, JET)	27
3.4.3	Protocole de signalisation OBS avec acquittement (TAW)	29
3.4.4	Protocole INI (Intermediate Node Initiated)	30
3.5	Résolution des contentions dans un réseau OBS-JET	31
3.5.1	Définition d'une contention	31
3.5.2	Domaine spatial : routage par déflexion	32
3.5.3	Domaine spectral : la conversion de longueurs d'onde	33
3.5.4	Domaine temporel : les lignes à retardement (FDL)	34
3.5.5	La segmentation	36
3.6	Détermination de l'offset time	36
3.6.1	Calcul de l'offset time (OT)	37
3.6.2	Effet de l'Extra Offset Time (EOT)	38
3.6.3	Taille fixe	38
3.6.4	Priorité des trafics	38
3.6.5	Utilisation de jetons aléatoires	39

3.6.6	Envoi agressif des paquets de contrôle	39
3.7	Initialisation des tables de routage	39
3.8	Prise en charge de la qualité de service	40
3.8.1	Utilisation de l'offset time	41
3.8.2	Utilisation de la segmentation	42
3.8.3	Perte anticipée	42
4	Accès au réseau OBS	44
4.1	Choix de la longueur d'onde	44
4.1.1	Max-SV	45
4.1.2	La longueur d'onde la moins utilisée sur le lien (LLWoL)	45
4.1.3	La longueur d'onde la moins utilisée par le noeud (LLWoLN)	45
4.1.4	La longueur d'onde la moins utilisée par le noeud voisin (LLWoNN)	46
4.2	Choix de l'offset time	46
4.2.1	Stratégie utilisant le chemin par défaut (S1)	46
4.2.2	Stratégie utilisant la taille moyenne des chemins (S2)	47
4.2.3	Stratégie purement adaptative (S3)	47
4.2.4	Choix de la date d'émission de la rafale	49
5	Routage en mode OBS	50
5.1	Routage par déflexion	50
5.1.1	Algorithme de routage	50

5.1.2	Algorithme de déflexion	50
5.1.3	Métriques	51
5.2	Initialisation des tables de routage	53
5.2.1	Motivations	53
5.2.2	Chemin le plus court	53
5.2.3	Modélisation du problème de flot	53
5.2.4	Utilisation de la solution du problème de flot	56
6	Simulation et résultats	58
6.1	Description des expériences	58
6.1.1	Topologie du réseau	58
6.1.2	Les modèles de flots utilisés	59
6.1.3	Les caractéristiques observées	60
6.2	Expérimentation des stratégies de choix de longueurs d'onde	60
6.2.1	Description	60
6.2.2	Commentaires sur les stratégies de choix de longueurs d'onde	61
6.3	Expérimentation des stratégies de déflexion	62
6.3.1	Description	62
6.3.2	Commentaires sur les stratégies de déflexion	62
6.4	Expérimentation des politiques de calcul de l'offset time	63
6.4.1	Influence des paramètres	63

6.4.2	Expérience 1	68
6.4.3	Expérience 2	69
6.4.4	Conclusion sur les politiques d'offset time	71
6.5	Initialisation des tables de routage	71
6.5.1	Description	71
6.5.2	Analyse des résultats	73
7	Conclusion	75

Liste des tableaux

2.1	Comparaison multicritère des modes de commutation optique	17
-----	---	----

Liste des figures

2.1	Fonctionnement de commutateurs à base de MEMS à deux dimensions . . .	8
2.2	Fonctionnement de commutateurs à base de MEMS à trois dimensions . . .	9
2.3	Fonctionnement d'un commutateur SOA	10
2.4	Architecture d'un réseau IP sur WDM	12
2.5	La signalisation en mode OBS	16
2.6	Comparaison des trois modes de commutation optique	18
3.1	Module électro-optique	20
3.2	Module tout optique	23
3.3	La signalisation sans acquittement	28
3.4	Deux exemples de signalisation sans acquittement	29
3.5	La signalisation avec acquittement TAW	30
3.6	La signalisation avec acquittement INI	31
3.7	Illustration du problème de contention	31
3.8	Problème d'OT insuffisant	33
3.9	Illustration du phénomène de fragmentation	35

3.10	Résolution du IOT par l'utilisation de FDL	36
3.11	Illustration de l'isolation d'une classe de service par EOT	41
6.1	Le réseau EON	59
6.2	Taux de perte pour les stratégies de choix de longueur d'onde	61
6.3	Taux de perte pour les stratégies de déflexion	62
6.4	Influence du paramètre γ sur le taux de perte	64
6.5	Influence du paramètre γ sur le délai	65
6.6	trafic variable utilisé pour l'étude du paramètre β	66
6.7	Influence du paramètre β	66
6.8	Les stratégies d'offset time	68
6.9	Evolution du taux de perte par IOT et de l'offset time aux perturbations de trafic	70
6.10	Délai et taux de perte avec un routage par la destination	72
6.11	Délai et taux de perte avec un routage par la source	72
6.12	Comparaison du routage par la source et du routage par la destination .	73

Liste des algorithmes

1	Algorithme de routage	51
2	Algorithme de déflexion	52
3	Algorithme itératif pour l'utilisation de l'objectif z_2	56

Liste des sigles et abréviations

Acronyme	Description	Première utilisation
DIR	Destination initiated Reservation	26
EDS	Early Dropping by Span	43
EDT	Early Dropping by Threshold	43
EOT	Extra Offset Time	37
FDL	Fiber Delay Line	35
GRWA	Groomed Routing Wavelength Assignment	15
IETF	Internet Engineering Task Force	15
INI	Initié par un Noeud Intermédiaire	30
IOT	Insufficient Offset Time	32
IP	Internet Protocol	11
IPOWDM	IP over WDM	11
JET	Just Enough Time	28
JIT	Just In Time	28
LACS	Link Capacity Adjustment Scheme	15
MEMS	MicroElectroMechanical Systems	7
MSPP	MultiService Provisioning Platform	15
OBS	Optical Burst Switching	2
OCS	Optical Circuit Switching	12
OPS	Optical Packet Switching	11
OT	Offset Time	3
REL	Realease message	27
RWA	Routing Wavelength Assignment	13
SIR	Source Initiated Reservation	26
SOA	Semi Optical Amplifier	8
TAW	Tell And Wait	29
TT	Temps de Traitement du paquet de contrôle	37
TTL	Time To Live	33
VCAT	Virtual Concatenation	15
WDM	Wavelength Division Multiplexing	10

Remerciements

Je remercie Brigitte Jaumard pour son soutien financier et humain, ainsi que pour la qualité de son encadrement qui représente une véritable introduction dans le monde de la recherche et dans le milieu académique.

Je remercie aussi ma co-directrice Halima Elbiaze, instigatrice du projet, pour son soutien financier.

Enfin je salue tous les membres de l'équipe dirigée par Brigitte Jaumard pour l'esprit de franche camaraderie et d'entraide qu'ils entretiennent tous les jours.

Dédicace

A ma maman.

Chapitre 1

Introduction

L'augmentation du nombre d'utilisateurs, l'omniprésence de l'Internet dans la nouvelle économie ainsi que l'explosion d'applications temps réel comme la téléphonie, la vidéo conférence ou encore la télévision contribuent toutes ensemble à un besoin grandissant en termes de bande passante. L'intégration des applications temps réel avec les services existants (web, transfert de fichiers ...) nécessite une attention particulière et la mise en place de mécanismes eux aussi coûteux en ressources. La fibre optique a démontré un large potentiel susceptible de subvenir aux besoins grandissants en bande passante. Cependant, la croissance ne s'arrêtera pas là : toujours plus d'utilisateurs désireront toujours plus de qualité et plus de services à moindre coût. A ceci s'ajoute la dérégulation des télécommunications qui génère une concurrence croissante et donc une politique de réduction des coûts.

Pour subvenir à cette demande, un redimensionnement consistant à augmenter les infrastructures sera toujours possible mais coûteux. On assiste donc à un engouement considérable de la recherche pour utiliser au mieux les ressources offertes par le médium optique. Cette optimisation s'opère à plusieurs niveaux. Tout d'abord sur la couche physique puisque l'amélioration des émetteurs et récepteurs laser permet, grâce au multiplexage spectral (WDM) d'augmenter les débits, mais aussi sur toutes les couches inter-

médiaires, jusqu'à la couche applicative où les progrès en matière d'encodage permettent, à qualité égale, de diminuer le besoin de bande passante.

Dans notre étude, nous nous intéressons aux protocoles d'accès au lien et de routage, c'est-à-dire que nous recherchons comment envoyer les données et les acheminer pour tirer au mieux parti du potentiel optique. Cette approche se confronte à une contrainte physique ou technologique inhérente au support optique : l'absence de mémoire. De ce fait, l'utilisation du *store and forward* est proscrite. La conception de nouveaux paradigmes de routage est nécessaire.

Pour relever le défi, une technique très prometteuse a déjà été proposée : la commutation optique par rafales (*Optical Burst Switching*, OBS). Il s'agit d'une adaptation de la commutation par rafales proposée dans les années 1980 par [1]. La version optique a été proposée par [2] en 1999. Cette technique, située entre la commutation de circuits et la commutation de paquets, consiste à envoyer de gros paquets de données (rafales) de l'ordre du méga-octet qui seront routés dans le domaine optique. Un paquet de contrôle est envoyé préalablement afin de réserver le chemin. Celui-ci contient les informations relatives à la rafale qui lui est associée et subit des conversions opto-electro-optique à chaque routeur afin que celui-ci réserve un lien sortant et une longueur d'onde pour la rafale. Le paquet de contrôle est donc ralenti à chaque noeud alors que les données transitent sans conversion électrique et donc sans délai à chacun des noeuds. Ainsi, en mode OBS, alors que les données sont encore à l'entrée du réseau, la signalisation et la réservation des ressources commencent à s'effectuer. De ce fait, aucune capacité de stockage des données n'est nécessaire à l'intérieur du réseau.

Les deux principaux aspects à améliorer dans les réseaux OBS résident dans le calcul du temps laissé au paquet de contrôle pour la signalisation et la réduction du taux de contention.

1.1 Problématique

Le principal défi pour la conception des réseaux optiques à commutation par rafales est la résolution des contentions. Une contention est observée lorsque deux rafales veulent emprunter la même longueur d'onde sur un même lien en même temps. En effet, pour réduire le délai des données, la rafale n'attend pas d'accusé de réception et rien ne peut garantir le bon cheminement de celle-ci.

Divers mécanismes ont été proposés dans la littérature pour permettre de résoudre les contentions. Certaines nécessitent des équipements coûteux et qui n'ont pas toujours atteint leur maturité.

Nous nous intéresserons à la déflexion qui peut être implantée sur des équipements standards et donc à moindre coût. Cette technique consiste à dévier une des rafales mise en jeu dans la contention sur un chemin alternatif.

Mais alors, nous ne pouvons connaître a priori la taille du chemin emprunté par la rafale pour arriver à destination. Comme l'entête est retardé à chaque noeud, si le chemin est plus long que prévu, l'entête peut être rattrapé par la rafale qui sera alors perdue. Ce problème est d'autant plus difficile à résoudre que le choix de l'*offset time*(OT) doit être fait à l'envoi des données, c'est-à-dire à l'accès.

Affecter un grand *offset time* permet d'assurer à l'entête de ne pas être rattrapé mais peut entraîner une sur-utilisation des ressources et augmenter la congestion. De plus, en minimisant l'*offset time* le délai à l'accès des données est réduit. Un mécanisme permettant de calculer l'*offset time* en fonction de l'état du réseau doit donc être mis au point pour pouvoir utiliser la déflexion.

1.2 Contributions

Dans ce projet, nous nous sommes penchés sur plusieurs aspects critiques pour la conception de réseaux OBS. Nous nous plaçons dans un réseau économique, c'est-à-dire que seule la déflexion sera utilisée pour résoudre les contentions. Dans un tel contexte, le calcul de l'offset time doit tenir compte de la congestion du réseau. Nous avons étudié plusieurs problèmes apparaissant à l'accès et à l'intérieur du réseau :

- le calcul de l'OT,
- le choix du lien de déflexion,
- le choix de la longueur d'onde à l'accès,
- l'initialisation des tables de routage avec un modèle de répartition de la charge.

Tous ces aspects sont confrontés lors de simulations réalisées sur le réseau optique européen EONET du projet cost_266 ([3]).

Les résultats obtenus sont les suivants :

- la répartition de la charge permet d'améliorer à la fois le taux de perte et le délai par rapport à une initialisation basée sur la topologie uniquement,
- le choix adaptatif de l'OT permet d'utiliser la déflexion, notamment en présence de trafic variable et améliore le compromis taux de perte/délai.

Certains de ces travaux ont déjà été publiés dans des conférences ([4],[5],[6]).

Chapitre 2

Architecture d'un réseau optique

2.1 Motivations pour l'utilisation de la lumière comme support d'information

Le médium optique s'impose comme la solution au transport massif d'informations. Il offre en effet plusieurs caractéristiques séduisantes non seulement en termes de bande passante, mais aussi sur le plan du signal et de l'installation. De plus, les ressources pour la fabrication des fibres semblent infinies. Pour plus de détail sur la couche physique, le lecteur peut se référer au premier chapitre de [7].

- **Largeur de bande.** Le multiplexage spectral peut être utilisé pour différentes sessions ou encore pour augmenter le débit d'une session. Certes, la couche physique devra porter une attention particulière à la forme du signal émis, à la précision des émetteurs et à la sensibilité des récepteurs. Les progrès physiques permettront d'élargir la plage de fréquence utilisable et de diminuer la taille de chaque longueur d'onde pour en envoyer de plus en plus simultanément. Dans les années 90, les premiers systèmes optiques à grande capacité (OC-48 permet d'atteindre 2.5Gbit/s) virent le jour et depuis, la capacité des liens augmente encore et les industriels annoncent déjà des débits de l'ordre du Tbit/s. Aujourd'hui, le OC-192

a été atteint pour un débit de 10Gbps.

- **La fiabilité.** Le signal lumineux dans la fibre optique subit moins d'atténuation que le signal électrique. Les photons n'exercent aucune influence entre eux, contrairement aux électrons qui se perturbent mutuellement et subissent des interférences électro-magnétiques. Les fibres optiques sont donc plus fiables que les câbles en cuivre.
- **Robustesse des liens.** La robustesse des fibres optiques permet de les installer dans toutes les conditions (utile dans un contexte industriel par exemple ou dans un contexte climatique rude).

La principale contrainte imposée par l'utilisation du médium optique repose sur l'absence de mémoire optique. Ainsi, si l'on veut utiliser des protocoles de commutation classiques, comme nous ne pouvons pas stocker la lumière en tant que telle, le seul moyen consiste à effectuer des conversions opto-électro-optique et utiliser des mémoires de type RAM. Cette alternative n'est pas envisageable de manière systématique, car elle retarderait trop les informations à transmettre et le gain dû à l'utilisation de la fibre serait amoindri.

Par la suite, nous présenterons les paradigmes alternatifs aux méthodes classiques de réseaux cuivrés, mais penchons nous tout d'abord sur le matériel optique.

2.2 Les fibres optiques

La conception des fibres est primordiale pour la qualité de transmission. Jue et Vokkarane expliquent que les paramètres les plus critiques sont la dispersion et l'atténuation ([8], pages 18 à 20).

Plus l'atténuation est importante, plus le signal devra être répété fréquemment ce qui impose un coût matériel. L'atténuation est due à des impuretés dans la fibre ou à des faibles variations dans l'indice de réfraction. Des amplificateurs peuvent être utilisés pour compenser la perte de puissance due à l'atténuation.

La dispersion quant à elle est principalement due au fait que différentes longueurs d'onde

ne se déplaceront pas toutes à la même vitesse. Ainsi, le signal s'étalera dans le domaine temporel. Cela est d'autant plus gênant si le codage correspond à la présence ou à l'absence de lumière : l'étalement dans le domaine temporel pourra fausser la valeur interprétée d'un bit (un 0 interprété comme un 1).

2.2.1 La fibre multimode

Le coeur de la fibre a un diamètre compris entre 50 et 100 microns. On distingue deux catégories.

- **La fibre à saut d'indice** : le coeur et la gaine optique sont composés de verre ce qui provoque de la dispersion. Elle permet d'atteindre des débits de l'ordre de 50 Mbps.
- **La fibre à gradient d'indice** : elle permet de réduire la dispersion constatée avec les fibres à saut d'indice. Pour ce faire, la gaine optique est composée de plusieurs couches de verre. Cette modification permet d'atteindre des débits de l'ordre de 1 Gbps.

2.2.2 La fibre monomode

Avec ce type de fibre, le coeur ne mesure plus que 10 microns de diamètre. L'effet direct est la diminution de la dispersion. Cependant, elle nécessite plus de puissance à l'émission.

2.3 Les commutateurs optiques

Une comparaison des différentes technologies en matière de commutateurs est proposée dans [9]. Nous nous pencherons plus particulièrement sur deux d'entre elles, à savoir les MEMSS (*MicroElectroMechanical Systems*) qui sont utilisées dans la technolo-

gie actuelle et les SOAs (*Semiconductor Optical Amplifier*) qui semblent être la technologie la plus pertinente pour une commutation à granularité fine (paquets ou rafales).

2.3.1 Les MEMS (*MicroElectroMechanical Systems*)

Les MEMS sont composés d'un ou de deux plateaux de miroirs dont le rôle est de dévier le signal du port d'entrée vers le port de sortie.

En deux dimensions (un plateau), le nombre de miroirs nécessaires est N^2 , avec N le nombre de ports d'entrée. Cette technologie est déjà déployée pour la commutation de circuits. Le fonctionnement de tels commutateurs est illustré par la figure 2.1.

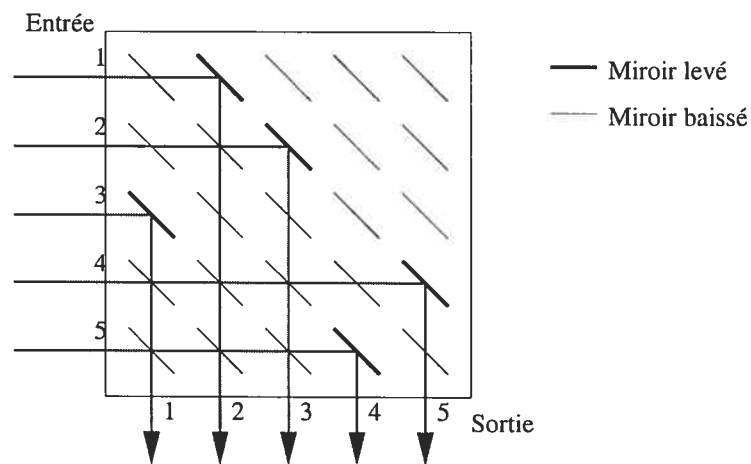


FIG. 2.1 – Fonctionnement de commutateurs à base de MEMS à deux dimensions

Une alternative, pas encore commercialisée de nos jours, consiste à utiliser deux plateaux de N miroirs chacun. Ainsi, le nombre de miroirs est diminué. Le fonctionnement des MEMS en trois dimensions est illustré par la figure 2.2.

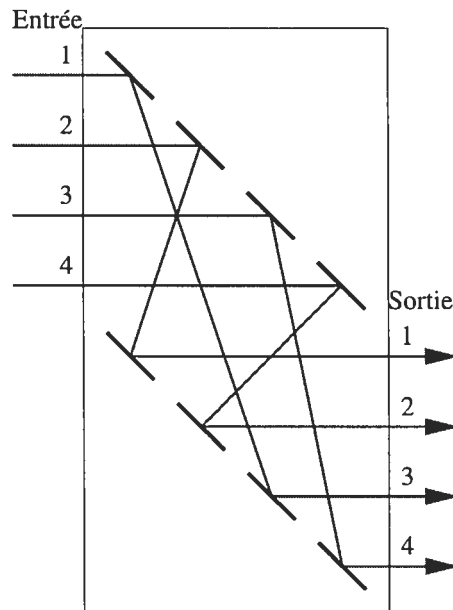


FIG. 2.2 – Fonctionnement de commutateurs à base de MEMS à trois dimensions

L'avantage des MEMS réside dans leur maturité pour une commutation tout optique. Cependant, ils souffrent de deux limitations :

- le nombre maximal de ports est encore limité à ce jour ;
- le temps de changement de configuration est très long (de l'ordre de 50 ms).

Le temps de changement de configuration de l'ordre de plusieurs millisecondes ([9]) ne permet pas d'utiliser cette technologie avec les paradigmes à fine granularité.

2.3.2 Les SOA (*Semiconductor Optical Amplifier*)

Les SOAs permettront bientôt de commuter les signaux en un temps de l'ordre de la nanoseconde ce qui est très satisfaisant pour la commutation de rafales et de paquets. Le fonctionnement est illustré par la figure 2.3. Le signal entrant est diffusé vers plusieurs SOAs grâce à un coupleur. La perte de puissance est ensuite compensée par le SOA. Une configuration correspond à un ensemble de SOAs allumés. Chaque SOA est relié à un coupleur dédié à un port de sortie. Les SOAs permettent donc de commuter très rapidement tout en offrant une diffusion simultanée vers plusieurs ports (*multicast*). Le

principal inconvénient est la perte de puissance du signal qui réduit la distance qu'il est possible de parcourir.

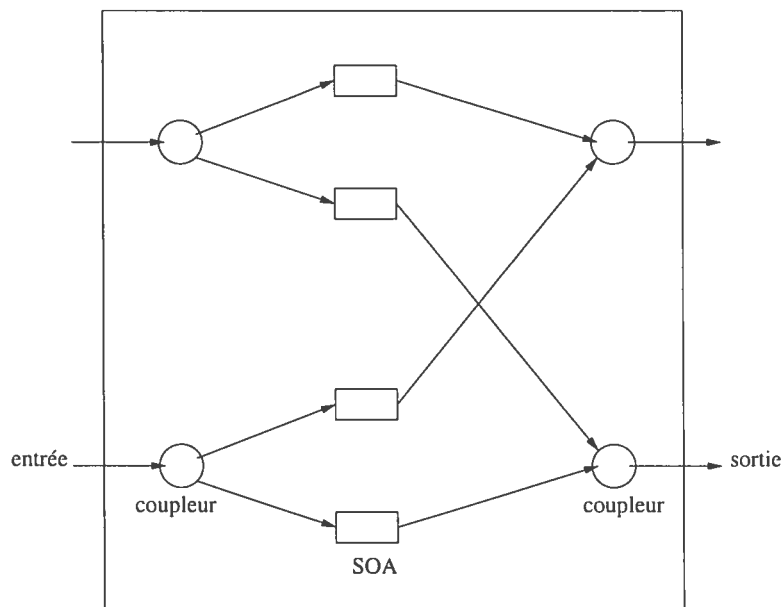


FIG. 2.3 – Fonctionnement d'un commutateur SOA

2.4 WDM

Si la conception des fibres optiques peut améliorer la capacité de celles-ci, le multiplexage permet de multiplier les capacités. Avec une largeur de bande de l'ordre de 25000Ghz, le médium optique offre un énorme potentiel de multiplexage spectral.

Le multiplexage spectral consiste à envoyer plusieurs signaux optiques simultanément sur la même fibre. Chacun de ces signaux emprunte un canal défini par une plage de fréquence. La dispersion du signal impose l'usage de plage de sécurité pour s'assurer que les signaux ne se chevauchent pas et que les récepteurs puissent différencier les canaux. Dès lors, pour envoyer le maximum de longueurs d'ondes dans une fibre, il faut que la plage associée soit minimum, de même que la déformation du signal. Il faut donc, au

niveau physique, travailler sur la précision des émetteurs et des récepteurs.

2.5 Caractéristiques actuelles des transmissions optiques

De nos jours, les capacités de transport les plus utilisées sont des OC-48 (2.5 Gbps) sur 40 longueurs d'ondes, mais il existe aussi des systèmes multiplexant 160 canaux OC-192 (10 Gbps), soit une capacité totale de 1.6 Tbps et, dans les laboratoires, des canaux à 80 Gbps offrant avec la technologie WDM une capacité totale de 12.8 Tbps. Alcatel a mis au point un prototype de commutateur OBS/OPS présenté dans [10].

2.6 Réseau IPoWDM

Nous considérerons les réseaux IPoWDM comme des réseaux d'accès constitués de liens optiques et de noeuds capables de router en tout optique (SOA). Par tout optique, nous entendons que le SOA ne nécessitera pas de conversion dans le domaine électrique pour commuter le signal. En plus, les commutateurs peuvent ajouter (ou retirer) du trafic provenant (ou à destination) de clients IP. Ces noeuds disposent donc d'interfaces entre le monde optique et le monde électrique.

Un tel réseau est illustré par la figure 2.4.

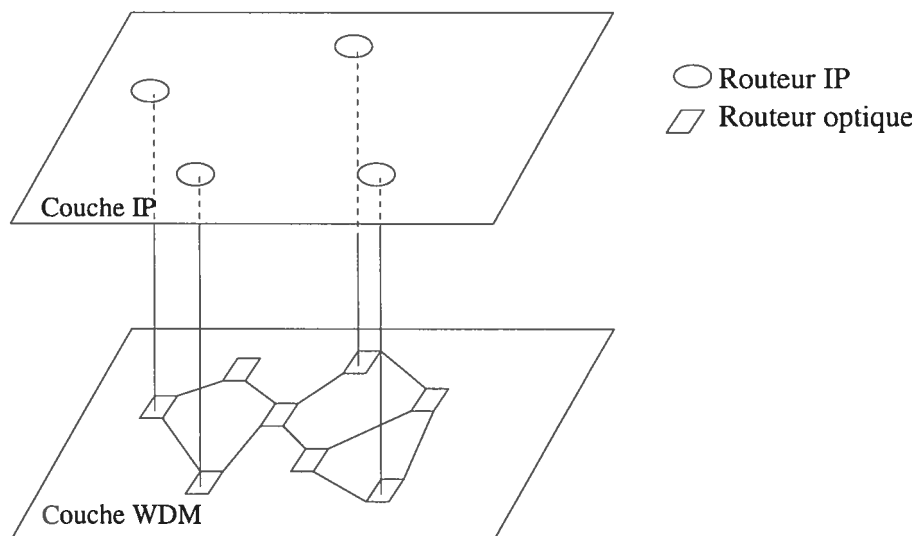


FIG. 2.4 – Architecture d'un réseau IP sur WDM

2.7 La commutation dans un réseau optique

2.7.1 Introduction

L'absence de tampons optiques nous oblige à reconsidérer une sous couche de la couche réseau. En particulier, nous devons reconsidérer les protocoles de commutation et de signalisation qui, dans le domaine électrique utilisent le *store and forward* comme réponse aux problèmes de contentions et de régulation de trafic. La littérature propose trois paradigmes pour la commutation dans les réseaux optiques :

- la commutation de circuits (OCS), orientée connexions ;
- la commutation de paquets (OPS), orientée paquets ;
- la commutation de rafales (OBS), intermédiaire entre circuits et paquets.

2.7.2 Commutation de paquets optiques (OPS)

La commutation de paquets optiques consiste à faire transiter les données de proche en proche sous forme de paquets. Un paquet est constitué d'un volume de données (de

l'ordre du mégaoctet), précédé d'un entête contenant les informations nécessaires pour le routage. On se rapproche beaucoup de IP, mais la contrainte imposée par l'absence de mémoire optique impose des solutions alternatives au *store and forward* utilisé dans IP.

L'OPS offre une bonne granularité, ce qui confère une bonne faculté d'adaptation aux variations de trafic et une bonne robustesse aux pannes. En plus, l'émetteur peut envoyer ses données dès qu'il le souhaite, sans négociation préalable, laissant aux autres maillons la charge de les router à destination.

Le principal frein au développement de la commutation de paquets optiques réside dans le fait que l'entête doit être interprété dans le domaine électrique pour apporter les informations nécessaires au routeur pour le routage du paquet associé. Ainsi, le paquet doit être retardé pendant la conversion O/E/O de son entête et le processus de routage. Cette approche nécessite donc des techniques de stockage, comme des fibres à retardement, qui n'ont pas encore atteint leur maturité ou encore l'utilisation de conversions opto-electro-optiques qui dégradent les performances en termes de délai.

2.7.3 Commutation de circuits optiques (OCS)

Dans un mode commuté par circuits, une connexion est établie pour toute la durée de la communication, comme dans le réseau commuté téléphonique.

Pour mettre en place une telle commutation, les tâches suivantes doivent être élaborées.

- La découverte de la topologie et des ressources disponibles peut être assurée par OSPF (*Open Shortest Path First*, [11]).
- Le routage et l'allocation de longueurs d'onde : connu sous le nom de RWA (*Routing and Wavelength Assignment*), ce problème a déjà été très étudié [12]). La modélisation évolue avec l'avancée des équipements et les contraintes varieront en fonction de l'équipement utilisé. Les contraintes fondamentales de ces formulations mathématiques sont inhérentes au support optique et sont décrites ci-dessous.

- La contrainte de conservation des flots : la proportion d'un flot f entrant dans un noeud est égale à la quantité de ce flot sortant sur ce noeud. Sauf à la source où la totalité du flot sort (sans entrer) et à la destination où la totalité du flot entre (sans sortir).
- La contrainte de continuité de longueur d'onde : sans conversion de longueur d'onde, le signal doit emprunter la même longueur d'onde de la source à la destination.
- La contrainte d'utilisation d'une longueur d'onde : une même longueur d'onde peut être utilisée par une et une seule connexion.
- La contrainte de capacité d'une longueur d'onde : la quantité totale de trafic affectée à une longueur d'onde ne doit jamais dépasser la capacité de celle-ci.
- La signalisation et la réservation des ressources : la signalisation est importante car elle assure la réservation du chemin jusqu'à la destination. Une fois les longueurs d'onde réservées, la transmission est généralement tout optique et ne nécessite ni temps de calcul ni capacité de stockage. En pratique cependant, les opérateurs misent sur le surdimensionnement du réseau, ce qui permet de planifier les connexions pour une longue période (i.e., de l'ordre du mois aujourd'hui).

Une fois les connexions établies, le trafic peut être envoyé avec une garantie ferme sur les critères de qualité de service. En effet, les contentions sont évitées par un contrôle d'admission facilité à la réservation du chemin et le délai sera garanti par le fait que les données ne subissent pas de stockage aux noeuds et utilisent, pour chaque connexion, toujours le même chemin. Cette architecture de réseau est aussi très simple à mettre en place tant que le réseau est surdimensionné. Ensuite, en augmentant la complexité des noeuds, on peut augmenter la charge du réseau.

Le mode de réservation avec accusé de réception impose un temps de latence pendant l'établissement des connexions. Ce temps de latence rend le paradigme peu robuste aux pannes. En effet, si un lien brise, il faut renégocier un nouveau chemin et les données devront attendre à l'accès pendant ce temps. Une solution aux problèmes de panne réside dans l'allocation de chemins de protection. Ils peuvent consister à protéger toute

une connexion ou encore juste un lien ou toute autre granularité intermédiaire. Dans tous les cas, l'utilisation de chemins de protection nécessite l'usage de capacité qui ne sera utilisée qu'en cas de bris de liens. De ce fait, les chemins de protection imposent une sous utilisation des ressources.

En présence de trafic variable, une partie de la capacité des liens sera gaspillée. Par exemple, pour une connexion transportant de la voix, les codecs imposent un débit, donc une certaine largeur de bande. Cette largeur sera réservée pour toute la connexion mais pas toujours utilisée (en particulier pendant les moments de silence).

De la même manière, des trafics hautement dynamiques augmentent la *burstiness* des connexions.

Les solutions permettant d'améliorer l'utilisation des ressources passent par l'élaboration de commutateurs de plus en plus puissants et complexes. Ainsi, les commutateurs dotés de module ADD/DROP pourront utiliser le groupage de trafic pour augmenter l'utilisation des liens (GRWA [13]). En effet, étant donné la capacité d'une longueur d'onde, la contrainte d'utilisation est un fort handicap pour une utilisation efficace des ressources du fait qu'une connexion remplira rarement une telle capacité. La contrainte d'utilisation est ainsi supprimée du problème RWA. Il faut cependant souligner que de tels modules nécessitent l'emploi de MSPPs pour ajouter ou retirer du trafic et régénérer le signal [14]. La commutation n'est alors plus tout optique lorsque la fonctionnalité est utilisée.

L'utilisation conjointe de flots bifurqués permet de remplir les longueurs d'onde. Les flots bifurqués consistent à répartir les données d'une même connexion sur plusieurs chemins. Les protocoles VCAT et LACS, supportés par l'IETF permettent d'utiliser cette fonctionnalité ([15]). Enfin, la conversion de longueur d'onde permet d'écarter la contrainte de continuité de longueur d'onde, mais est cependant une solution encore très coûteuse.

2.7.4 Commutation de rafales (OBS)

L'idée de rafales introduite dans [1] consiste à réserver un lien pendant un certain temps et à envoyer plusieurs paquets durant tout ce temps, afin d'utiliser au mieux la bande passante réservée.

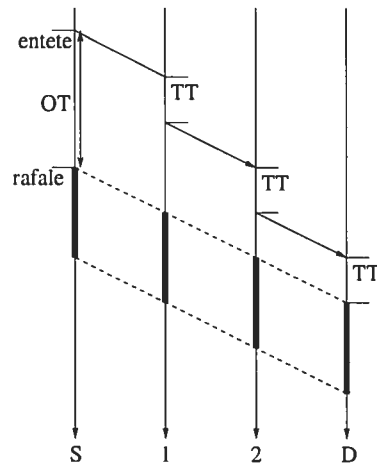


FIG. 2.5 – La signalisation en mode OBS

Le concept a été adapté aux réseaux optiques dans [2] pour mettre au point un mode de commutation proposant un compromis entre le temps de signalisation et le temps de calcul pour le routage. Il offre une bonne granularité (car le trafic d'une session est envoyé en plusieurs rafales routées indépendamment les unes des autres) et donc une bonne robustesse aux pannes et aux variations de trafic.

Cette solution est en fait très proche de l'OPS, mais le problème de mémoire inhérent à l'OPS est résolu par l'utilisation d'un protocole particulier de signalisation. Ainsi, la rafale sera routée dans le domaine optique.

Signalisation en mode OBS

L'objectif du protocole est de réserver un chemin en minimisant le temps de latence à l'accès et le temps de traitement à chaque noeud, tout en satisfaisant la contrainte imposée par l'absence de mémoire optique. L'approche orientée circuit élimine le temps de processus en introduisant un temps de latence lors de l'établissement de la connexion. La commutation de paquet élimine le temps de latence au prix d'un temps de calcul élevé et de la nécessité de mémoire.

En mode OBS, on utilisera un protocole unidirectionnel afin de réduire le temps de

latence avant l'envoi des données par rapport aux protocoles avec accusé de réception cependant plus sûrs.

La figure 2.5 illustre le mécanisme : les données sont envoyées de S à D via les noeuds 1 et 2.

Un paquet de contrôle précède la rafale d'un temps OT appelé *offset time*. Le paquet de contrôle contient toutes les informations sur la rafale qui le suit (notamment la date de début, la taille et la destination). Il est chargé de réserver un lien avant l'arrivée de la rafale pour que le commutateur puisse configurer le SOA et router dans le domaine optique (sans stockage, sans calcul et sans conversion O/E/O). Ainsi, une partie du OT sera consommée à chaque noeud en temps de traitement du paquet de contrôle (TT). Le gain de la technique par rapport à la commutation de circuits vient du fait que la rafale est envoyée avec un plus faible temps de latence par rapport à l'établissement d'une connexion. Ce temps de latence est censé permettre le traitement de l'entête préalablement à l'arrivée de la rafale pour chaque noeud. Cela résoud la principale difficulté de l'OPS puisque les lignes à retardements ne sont plus nécessaires pour faire patienter la rafale en attendant le choix du port de sortie. Un autre avantage de OBS sur OCS provient du fait que la granularité peut être très fine dépendamment du mode de réservation utilisé.

2.7.5 Conclusions

Commutation	Utilisation	Signalisation	Tampon	Adaptation	Calcul
OPS	BONNE	NON	OUI	BONNE	OUI
OCS	MAUVAISE	AVEC ACCUSE	NON	MAUVAISE	NON
OBS - JET	BONNE	SANS ACCUSE	NON	BONNE	OUI

TAB. 2.1 – Comparaison multicritère des modes de commutation optique

La figure 2.6 illustre le mode de fonctionnement de base des différents modes de commutation ; les données sont envoyées de S à D via trois noeuds intermédiaires 1, 2 et 3. La table 2.1 compare les caractéristiques des différents modes de commutation dans

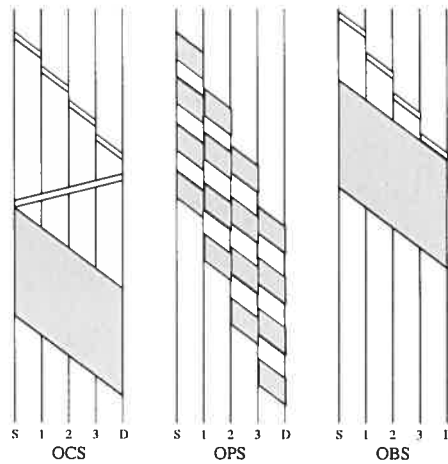


FIG. 2.6 – Comparaison des trois modes de commutation optique

les réseaux optiques en termes de l'utilisation de la bande passante, de la signalisation, de nécessité de mémoire optique ou électrique suite à des conversions O/E/O, de faculté d'adaptation aux pannes et aux variations de trafic et de temps de calcul aux noeuds. Il ressort de ces illustrations que le routage dans le domaine optique est un compromis entre performances, complexité et maturité du matériel. Si la commutation de circuit nécessite un matériel déjà disponible, c'est en partie parce que sa mise ne place est plus simple et donc que les constructeurs et fournisseurs ont tout d'abord opté pour elle. D'un autre côté, la recherche de performances pousse vers une meilleure granularité et un matériel pas encore mature.

L'OBS représente un compromis entre la complexité, le prix du matériel et la granularité.

Chapitre 3

Architecture des réseaux OBS

3.1 Le réseau

Nous considérons un réseau d'accès en mode OBS composé d'un coeur tout optique. Par tout optique, nous entendons que les routeurs ne peuvent pas faire de conversion opto-electro-optique sauf pour traiter les paquets de contrôle qui transitent sur une longueur d'onde dédiée. Ainsi, le trafic transitera dans le réseau sous forme de signal optique et les contentions seront résolues par déflexion.

Les noeuds d'accès capables d'effectuer les conversions de médium seront chargés de récupérer les paquets IP et de les convertir en signaux lumineux à destination d'un noeud tout optique (ADD). De la même manière, ils pourront convertir le signal lumineux pour découper les rafales arrivées et transmettre des paquets IP dans le réseau cuivré (DROP).

3.2 Les routeurs OBS

Les noeuds optiques permettent d'accumuler du trafic IP pour construire des rafales qui seront ensuite injectées dans le coeur optique (module électro-optique), mais ils

disposent aussi d'un module tout optique capable de router les rafales en transit sans mémoire ni conversion dans le domaine électrique.

3.2.1 Module électro-optique

Le module électro-optique a pour rôle de générer du trafic optique à partir de paquets IP. Il consiste donc à accumuler des paquets IP ayant des caractéristiques communes (destination, critères de qualité de service). Ces paquets seront ensuite concaténés en rafales qui seront envoyées dans le coeur du réseau sous forme de signaux lumineux. La figure 3.1 illustre l'architecture d'un module électro optique.

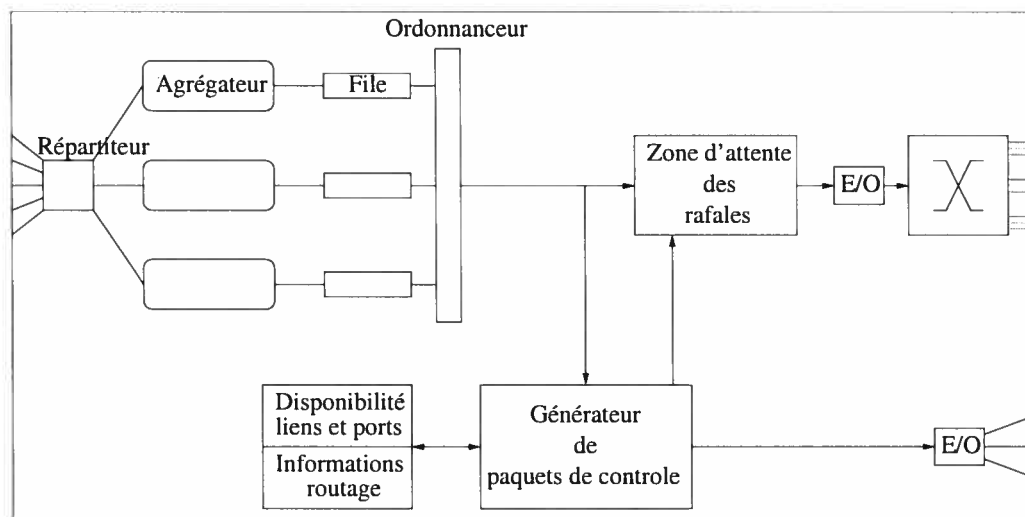


FIG. 3.1 – Module électro-optique

Considérons plus en détail chaque composante du module.

Le répartiteur

Le répartiteur reçoit des paquets IP puis les trie pour les envoyer vers les agrégateurs. Le tri suit la définition des classes d'équivalence. Typiquement, une classe d'équivalence sera caractérisée par le noeud optique de sortie et les critères de qualité de service.

Les agrégateurs

Les agrégateurs récupèrent des paquets IP triés par classe d'équivalence. Ces paquets IP sont concaténés jusqu'à l'obtention d'une rafale. Un agrégateur est caractérisé par la classe d'équivalence qu'il sert, la taille du conteneur (la rafale est prête lorsque le conteneur est plein) et un temps limite au bout duquel la rafale est envoyée quelle que soit sa taille. Une fois prête, la rafale est placée dans une file d'attente dédiée à chaque classe d'équivalence.

L'ordonnanceur

L'ordonnanceur choisit la classe d'équivalence à servir et envoie un signal au générateur de paquets de contrôle contenant la taille de la rafale et la classe d'équivalence.

Le générateur de paquets de contrôle

Connaissant la taille de la rafale, la date et la classe d'équivalence, le générateur de paquets de contrôle consulte une table contenant la disponibilité de chaque port et de chaque longueur d'onde. Il détermine alors le port de sortie, la longueur d'onde et la date de départ de la rafale (la date dépend de l'*offset time* et de la disponibilité du lien). Celui-ci est alors placé dans une file d'attente triée par date de départ des rafales. Le paquet de contrôle est converti en signal optique et envoyé sur la longueur d'onde

dédiée du lien choisi alors que les rafales ne sont converties et envoyées qu'à la date définie par le générateur de paquet de contrôle.

3.2.2 Module tout optique

Le module tout optique est utilisé pour les rafales en transit. La figure 3.2 représente l'architecture du module.

Traitement du paquet de contrôle

Le module dispose de ports dédiés aux paquets de contrôles. Ceux-ci sont convertis dans le domaine électrique afin de fournir les informations sur la rafale au calculateur de route. Celui-ci les récupère et consulte la table de disponibilité et les informations de routage pour exécuter l'algorithme de routage. Une fois terminé, le SOA est préprogrammé, le paquet de contrôle est mis à jour, reconverti en signal lumineux et envoyé sur le port choisi.

Le brasseur optique

Le brasseur optique dispose de ports optiques d'entrée et de sortie. Il est programmé par le module de routage pour relier un port d'entrée à un port de sortie durant une certaine période de temps correspondant au passage d'une rafale. Ainsi, les données transitent en restant dans le domaine optique. Nous ne tenons pas compte du temps nécessaire pour l'aiguillage (de l'ordre de la nanoseconde avec des commutateurs SOA). Sinon, il faudrait laisser autant de temps entre chaque rafale au lieu de les coller.

3.2.3 Informations pour le routage

Les tables de routage contiennent, pour chaque destination, un lien par défaut et, pour chaque lien et chaque destination, la taille du chemin le plus court ainsi qu'un indicateur de routage alternatif. Nous avons investigué plusieurs indicateurs et métriques qui seront présentées dans le chapitre suivant.

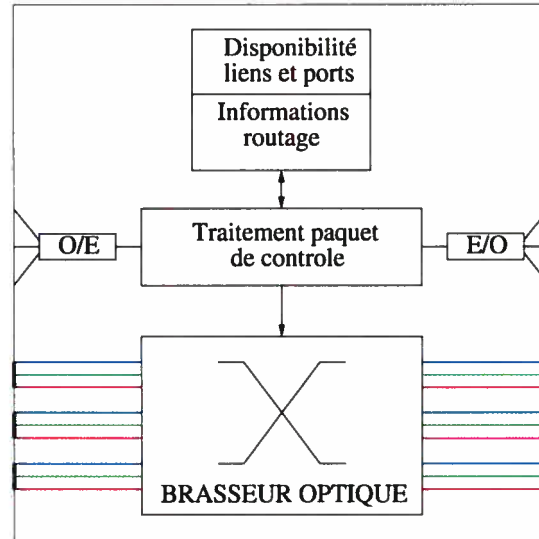


FIG. 3.2 – Module tout optique

3.3 Création des rafales

Dans notre projet, nous ne nous intéressons pas au processus d'agrégation en tant que tel et considérons que l'interarrivée des rafales et leur taille suivront une distribution exponentielle. Cependant nous respectons le fait que le conteneur est défini par une taille maximum et bornerons la taille des rafales (par exemple 1.5Mbps).

Le processus d'agrégation ne demeure pas moins une composante essentielle du OBS, si bien que nous présentons dans cette partie les bases de l'agrégation des rafales. Pour

plus d'information sur l'agrégation des rafales, le lecteur pourra se référer à [16].

3.3.1 Architecture du module de création de rafales

Le noeud d'accès dispose de ports d'entrée recevant des paquets IP. Ces paquets sont triés en plusieurs files caractérisées par une classe d'équivalence au sens de MPLS (destination et critères de qualité de service). Les paquets sont accumulés dans la file jusqu'à l'obtention d'une rafale. Ensuite, un paquet de contrôle correspondant à la rafale est créé puis envoyé, suivi par la rafale. La question qu'il reste à éclaircir est la définition d'une rafale : quand considère-t-on que la rafale est prête à être envoyée ?

3.3.2 Déclenchement d'une rafale

Basé sur le temps

Une première solution consiste à affecter un temps pour chaque classe d'équivalence. Dès que le temps est écoulé, le contenant est vidé et les paquets à l'intérieur forment la rafale. De cette manière, l'interarrivée des rafales est constante quelles que soient les variations du taux d'arrivée des paquets, ce qui engendre des tailles de rafales très variables.

Basé sur la taille

Une autre méthode consiste à attendre que le contenant soit plein. Ainsi, la taille des rafales sera constante, mais l'interarrivée variera avec la charge : si la charge diminue, le délai de création de la rafale sera allongé.

Mixte

La solution adoptée consiste à combiner les deux approches en considérant les deux éléments déclencheurs. Donc la rafale sera prête à être envoyée si le contenant est plein ou si le temps alloué est dépassé. Le mécanisme permet de borner l'interarrivée et la taille des rafales.

3.4 Signalisation dans un réseau OBS

La commutation OBS est initialement conçue pour une utilisation avec signalisation JET sans accusé de réception. Cependant, on trouve maintenant dans la littérature un petit nombre de contributions utilisant d'autres protocoles de signalisation, dont certains utilisent un accusé de réception ([17]).

3.4.1 Classification des protocoles de signalisation

Avec accusé de réception, sans accusé de réception ou hybride

Avec un protocole de signalisation sans accusé de réception, la source demandera la réservation des ressources sur le chemin, puis enverra la rafale sans attendre de réponse. Comme la réservation n'est pas acquittée, il se peut que des données soient perdues.

Avec un acquittement par contre, les données ne seront envoyées que si les ressources ont effectivement pu être réservées. Le problème est alors que les données, avant d'être envoyées devront attendre l'accusé de réception.

Cette caractéristique du protocole de signalisation jouera donc un rôle sur le compromis entre la perte et le délai d'accès.

Une technique hybride consiste à confirmer la réservation entre la source et un noeud intermédiaire, c'est-à-dire pour une partie du chemin. Le temps de latence à l'accès est

ainsi réduit et, si l'on ne peut garantir l'acheminement du paquet jusqu'à la destination, on assure toujours l'acheminement jusqu'à un certain noeud intermédiaire.

Initié par la source (SIR), par la destination (DIR) ou par un noeud intermédiaire

Si la réservation est initiée par la source (SIR), celle-ci envoie une requête qui, à chaque noeud de la route réservera les ressources jusqu'au noeud suivant, et ce jusqu'à la destination. Si le protocole impose un accusé de réception, alors la destination enverra la confirmation de la réservation et la rafale pourra être envoyée. La principale cause de blocage provient du manque d'information au niveau de la source.

Dans le cas d'une réservation initiée par la destination (DIR), la source enverra un entête jusqu'à la destination, celui-ci récoltera l'information nécessaire sur chaque lien du trajet et, une fois arrivée à destination, une longueur d'onde sera choisie et une requête de réservation envoyée vers la source. Cette requête réservera alors les ressources de noeud en noeud. Le blocage dans le cas d'une réservation initiée par la destination est dû à des informations obsolètes.

Pour ce critère, comme pour l'accusé de réception, le compromis consiste à initier la réservation d'un noeud intermédiaire. Ainsi, la réservation sera DIR entre la source et ce noeud et SIR entre le noeud et la destination.

Réservation immédiate ou retardée

Lorsque l'entête arrive à un noeud, doit-il réserver la ressource immédiatement ou alors seulement pour la date d'arrivée de la rafale ? La réservation immédiate est plus simple à implanter mais engendre une mauvaise utilisation des liens et un plus gros taux de blocage. La réservation retardée par contre permet de réduire le taux de blocage en utilisant plus efficacement les ressources. Le prix à payer est une plus grande complexité. Le paquet de contrôle doit en effet connaître le temps le séparant de la rafale. De plus,

la réservation d'un intervalle de temps crée des intervalles libres et les noeuds doivent donc conserver plus d'informations sur l'utilisation de chaque longueur d'onde.

Fin de la réservation implicite ou explicite

Une fin de réservation implicite est annoncée par l'entête qui doit alors connaître la taille de la rafale et le temps le séparant de celle-ci. En mode explicite, la réservation prend fin lorsque le noeud reçoit un message de fin (REL). Ce message est envoyé après la rafale. Mettre fin à la réservation de façon implicite permet d'améliorer l'utilisation de la bande passante du fait que le canal sera libéré immédiatement à la fin de la rafale alors que la libération explicite nécessite l'attente et le traitement d'un paquet.

Synthèse des catégories de signalisation

Les protocoles avec accusé de réception assurent la réservation au prix d'un délai additionnel par rapport aux protocoles sans accusé de réception. La réservation retardée permet d'utiliser la bande passante de façon plus efficace que la réservation immédiate, au prix d'une plus grande complexité. De façon similaire, la libération implicite offre une meilleure utilisation des ressources au prix d'une plus grande complexité que la libération explicite.

3.4.2 Protocoles de signalisation OBS sans acquittement (JIT, JET)

La figure 3.3 illustre le protocole sans acquittement.

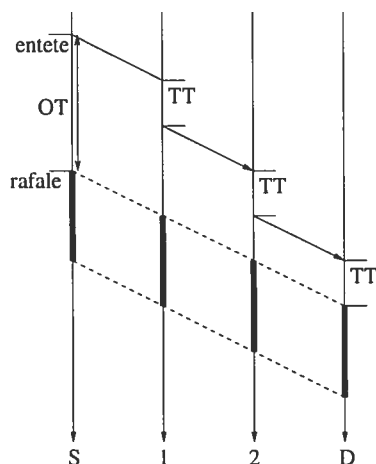


FIG. 3.3 – La signalisation sans acquittement

Just Enough Time (JET)

Le protocole JET (*Just Enough Time*) [2] est à réservation retardée et à libération implicite. De ce fait, le paquet de contrôle devra en permanence connaître l'*offset time* qui le sépare de la rafale ainsi que la taille de celle-ci. Ainsi, le SOA sera configuré à l'arrivée de la rafale et une autre rafale pourra la suivre pratiquement immédiatement (1 ns pour modifier la configuration avec des SOA).

Just In Time (JIT)

Le protocole JIT (*Just In Time*) est à réservation immédiate et à libération explicite [18]. Ainsi, le paquet de contrôle n'a besoin de connaître que la destination de la rafale pour être pris en charge. Le brasseur optique sera configuré dès l'algorithme de routage terminé et jusqu'à l'arrivée du paquet mettant fin à la réservation. La figure 3.3 illustre le protocole sans acquittement. La réservation immédiate et la libération explicite engendrent une mauvaise utilisation des ressources, mais facilitent l'implantation. La figure 3.4 permet de comparer JIT et JET. Sur cette figure, l'axe du temps est orienté de droite à gauche, ce qui permet de voir défiler les rafales de gauche à droite. Avec le protocole

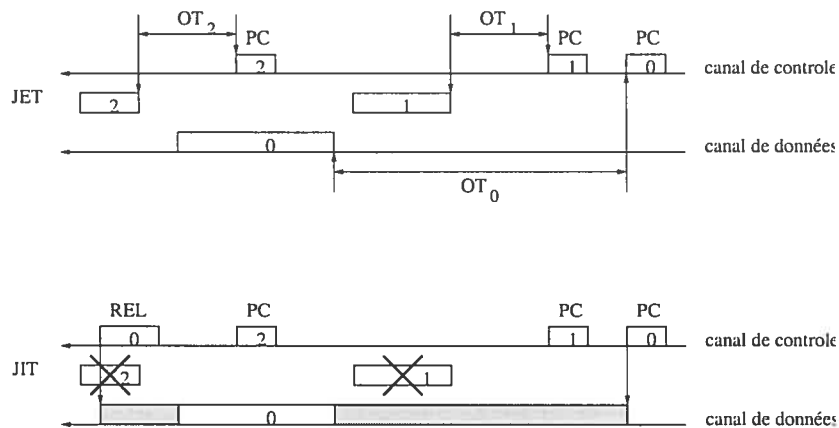


FIG. 3.4 – Deux exemples de signalisation sans acquittement

JET, les trois rafales peuvent être servies alors que JIT utilise trop de bande passante pour servir la rafale 0, si bien que les deux autres seront perdues.

3.4.3 Protocole de signalisation OBS avec acquittement (TAW)

Le protocole TAW (*Tell and Wait*) [8] est un protocole à réservation retardée avec libération explicite. Le fait d'attendre l'accusé de réception engendre un délai additionnel à l'accès par rapport aux protocoles sans accusé de réception. Cependant, il assure l'acheminement de la rafale une fois lancée. Il peut être initié soit par la source, soit par la destination. La figure 3.5 illustre la signalisation TAW.

Agustí-Torra, Bochmann et Cervelló-Pastor ([17]) proposent d'utiliser un mécanisme de retransmission de rafales en réponse aux contentions. Ainsi, si une contention apparaît, le noeud la constatant réservera un intervalle de temps libre et enverra une requête de retransmission à la source. Si la congestion devient trop importante, les auteurs proposent d'utiliser directement le TAW en mode DIR, c'est-à-dire qu'une requête est envoyée à la destination et les ressources sont réservées sur le chemin du retour. Dans

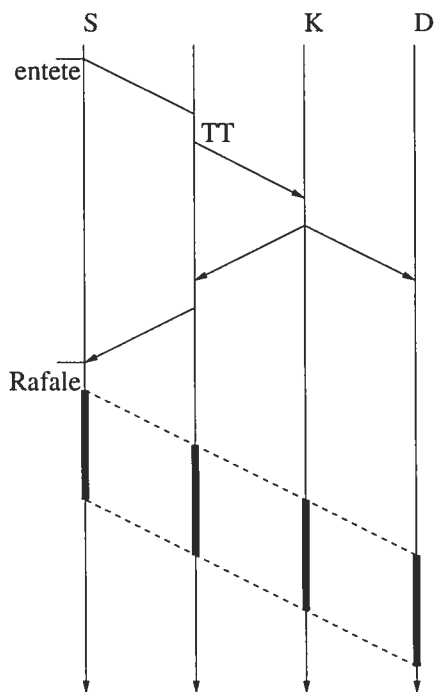


FIG. 3.6 – La signalisation avec acquittement INI

3.5 Résolution des contentions dans un réseau OBS-JET

3.5.1 Définition d'une contention

Une contention apparaît lorsqu'un paquet de contrôle tente de réserver une longueur d'onde déjà occupée sur un lien. La figure 3.7 permet d'illustrer le problème. Deux

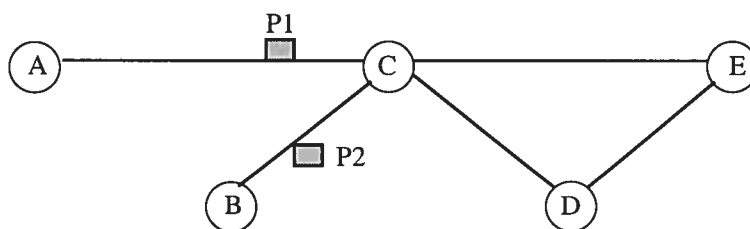


FIG. 3.7 – Illustration du problème de contention

paquets de contrôle $P1$ et $P2$ réservent les ressources pour des rafales voyageant sur la même longueur d'onde et destinées à emprunter le lien CE . En notant S_i la date à

laquelle le premier bit de la rafale signalé par la paquet de contrôle P_i emprunte le lien CE et E_i la date à laquelle le dernier bit emprunte CE , on constate une contention si :

$$\exists i \neq j \text{ tel que } S_i \leq S_j \leq E_i.$$

Sans tampon optique, l'utilisation du *store and forward* est proscrite et la résolution des contentions devient une question majeure. La littérature propose cependant trois techniques de résolution plus une approche permettant de réduire le volume mis en jeu par les contentions.

3.5.2 Domaine spatial : routage par déflexion

Définition du routage par déflexion

Le routage par déflexion consiste à réserver un lien alternatif pour une des rafales mise en jeu dans la contention.

Ainsi, dans l'illustration de la figure 3.7, le premier paquet de contrôle réservera le lien par défaut CE et le second sera défléchi vers le chemin CDE . Cette solution offre l'avantage de pouvoir être mise en place avec un matériel standard et donc à moindre coût.

Difficulté

Le chemin alternatif n'est pas forcément de la même longueur que le chemin par défaut. En particulier, si le chemin est plus long, le temps alloué pour le traitement du paquet de contrôle peut devenir insuffisant (IOT) auquel cas la rafale rattrapera le paquet de contrôle. Dès lors, la signalisation n'est plus assurée à l'arrivée de la rafale et celle-ci sera perdue (figure 3.8). Ce problème peut être résolu par l'emploi de mémoire à court terme ou par conversion O/E/O permettant l'usage du *store and forward*. Dans le cadre de notre étude, nous ne considérons pas la première alternative à cause du coût

de l'équipement et la seconde pour des raisons de performance.

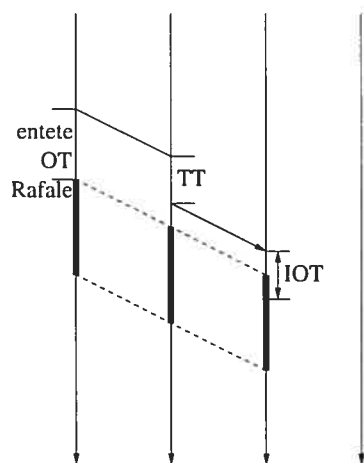


FIG. 3.8 – Problème d'OT insuffisant

Dès lors, l'attribut le plus important avant d'injecter un paquet dans le réseau sera l'OT qui correspond indirectement à une durée de vie (*Time to Live*, TTL). Il faudra alors mettre au point des mécanismes permettant d'évaluer la taille du chemin en considérant l'état de congestion du réseau. Une proposition sera présentée dans le chapitre suivant.

3.5.3 Domaine spectral : la conversion de longueurs d'onde

Définition de la conversion de longueurs d'onde

Le médium optique permet un large multiplexage, chaque lien proposant ainsi plusieurs longueurs d'onde. Le nombre de longueurs d'onde pouvant être différenciées sur une fibre dépend principalement de la précision des émetteurs et des récepteurs au niveau physique. Une solution pour la résolution des contentions consiste à régénérer la rafale mise en jeu par la contention sur une longueur d'onde alternative ([20]).

Difficulté

La principale difficulté provient du fait que les convertisseurs de longueurs d'onde ne sont pas réellement matures du point de vue physique. Ensuite, le coût du réseau augmentera avec l'utilisation de tels commutateurs. Enfin, la conversion implique un délai dans le temps de traitement des paquets de contrôle (choix de la longueur d'onde alternative) et dans la transmission du champ d'information utile (régénération du signal). Certaines solutions ont été proposées dans la littérature pour le choix de la longueur d'onde alternative, notamment par Xu, Qiao, Li, Xu dans [21]. Ils présentent tout d'abord HORIZON (rapide mais peu efficace) et LAUC-VF (efficace mais trop lent) avant de présenter une série d'autres politiques (Min-SV, Min-EV, Max-SV, Max-EV, Best-fit). Le principal critère à considérer est la fragmentation créée par la rafale : en insérant une rafale dans un intervalle, deux nouveaux intervalles sont créés. La figure 3.9 illustre les effets des politiques proposées dans [21] en termes de fragmentation :

- Max-SV consiste à choisir la longueur d'onde qui n'a pas été utilisée depuis le plus longtemps (on trouve parfois cette stratégie sous le nom de Lauc-VF),
- Min-SV au contraire choisira de coller la rafale à la rafale la plus proche, minimisant ainsi la taille du trou précédant la nouvelle rafale,
- Max-EV maximise la taille du trou suivant la rafale,
- Min-EV minimise la taille du trou suivant la rafale,
- Best-fit choisit d'insérer la rafale dans le plus petit intervalle libre.

3.5.4 Domaine temporel : les lignes à retardement (FDL)

Définition des lignes à retardement

Si l'on ne peut pas stocker la lumière en tant que telle, on peut toujours la retarder en augmentant sa distance à parcourir ([20]). On utilise pour cela des liens formant une boucle. Ils permettent de supprimer certaines contentions. En effet, si la rafale associée

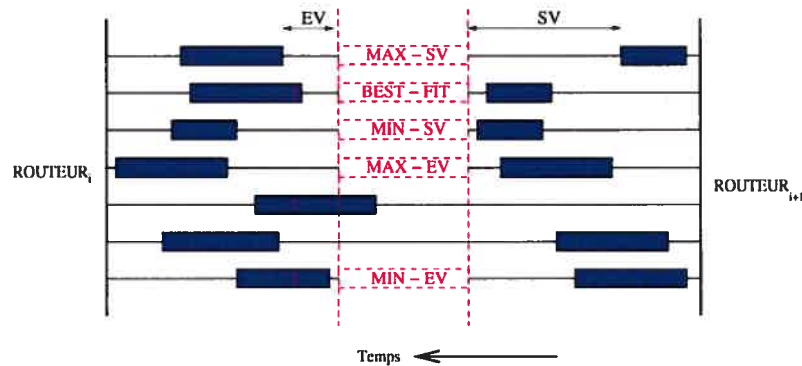


FIG. 3.9 – Illustration du phénomène de fragmentation

à P_i est en conflit avec celle associée à P_j , alors $S_i < S_j < E_i$. La contention disparaît si $S_j + T_{FDL} > E_i$ en notant T_{FDL} la durée de la fibre à retardement.

De plus, les FDL permettent de résoudre le problème de IOT qui peut apparaître lors du routage par déflexion en retardant la rafale (figure 3.10).

Problème

Les limites de cette alternative sont au niveau physique puisqu'à ce jour, les tampons de mémoire à court terme sont toujours du domaine de la recherche [22].

Le coût du réseau devrait croître grandement avec l'usage de tel matériel. Étant donné la vitesse de propagation du signal, la longueur des fibres à retardement devra être considérable (300km/ms).

Il faut souligner que cette solution ne représente pas une solution de stockage étant donné que lorsque l'on engage une rafale sur une FDL, on ne pourra pas la router avant qu'elle n'en sorte. Cela complexifie la conception des commutateurs car il faudra proposer plusieurs longueurs de fibres à retardement et le choix de la fibre devra faire l'objet d'une attention particulière.

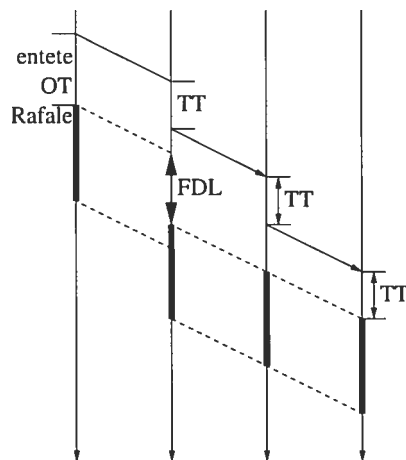


FIG. 3.10 – Résolution du IOT par l'utilisation de FDL

3.5.5 La segmentation

L'approche présentée dans [23] consiste à segmenter la rafale en contention. Le volume qui n'est pas mis en jeu par la contention sera traité normalement et seul le volume en conflit sera traité par une méthode alternative (déflexion, conversion ou FDL). L'approche est intéressante car les petites rafales auront plus de chance de trouver un chemin que les grosses. Cependant, la segmentation de la rafale impose un traitement électronique et retarde les données. Les résultats présentés dans [23] montrent que le volume de données perdu est diminué.

3.6 Détermination de l'offset time

La littérature ne couvre pas encore la question du calcul de l'OT. La majorité des travaux portant sur cette question sont effectués dans un but orienté vers la qualité de service mais plutôt qu'une métrique, ces travaux donnent un mécanisme permettant de calculer l'OT d'une classe de trafic relativement à d'autres classes.

3.6.1 Calcul de l'offset time (OT)

L'OT est consommé à chaque noeud. En effet, à chaque noeud, le commutateur est préprogrammé et la rafale n'est pas ralentie. Le paquet de contrôle quant à lui subit les traitements énumérés ci-dessous.

- Conversion opto-électrique : le paquet de contrôle détient toute l'information essentielle pour la signalisation et le routage. Il doit donc être traité par un processeur et la conversion opto-électrique est nécessaire.
- Délai de files d'attente : un même routeur recevra une multitude de paquets de contrôle et devra les traiter séquentiellement ce qui impose des délais de files d'attente.
- Processus : le port de sortie est déterminé par une procédure de routage qui demande un temps de processus.
- Conversion électro-optique : une fois la tâche accomplie, le PC est envoyé vers le prochain noeud en utilisant le médium optique.

Nous avons fait l'hypothèse simplificatrice communément admise que le temps de traitement est constant. Cela implique que le délai de file d'attente de la couche contrôle est considéré comme constant. Cela simplifie le problème puisque désormais, nous pouvons exprimer l'OT comme une fonction linéaire de la taille du chemin suivi. Soient TT le temps perdu par le paquet de contrôle à chaque noeud et N la taille du chemin en nombre de liens :

$$OT = N \times TT.$$

Si N_{SR} est la taille du chemin le plus court :

$$OT \geq N_{SR} \times TT.$$

La déflexion peut engendrer l'utilisation d'un chemin plus long que le chemin prédéfini. Ainsi, si $N \geq N_{SR}$, la rafale sera perdue par IOT. Nous exprimerons donc l'OT par

$$OT = (N_{SR} + EOT) \times TT,$$

où EOT représente l'allongement possible du chemin utilisé relativement au chemin le plus court.

3.6.2 Effet de l'Extra Offset Time (EOT)

Le choix de l'EOT est le principal élément de notre étude. Si l'OT représente indirectement le TTL de la rafale, il faudra l'augmenter pour offrir à la rafale un plus large choix de chemin. D'un autre côté, le paquet de contrôle est envoyé une fois la rafale prête et celle-ci attendra l'OT avant d'être envoyé à son tour. De ce fait, l'OT impose un temps de latence avant l'envoi de la rafale. Le compromis entre le délai à l'accès et le TTL sera donc primordial et il ne faudra pas utiliser de trop grands OT.

Un autre phénomène impose une attention particulière à l'OT. Etant donné deux flots, celui à qui l'on offre le plus grand OT se verra implicitement affecter une plus grande priorité, au détriment de l'autre ([24]). Ainsi, si nous ne voulons pas engendrer ce genre de phénomène, il faudra affecter les EOT de façon équitable.

Enfin, de trop grands OT permettent aux rafales d'utiliser de plus longs chemins, et donc d'utiliser plus de ressources dans le réseau. De ce fait, ils peuvent contribuer à amplifier le phénomène de congestion.

3.6.3 Taille fixe

La plupart des contributions ne traitant pas du calcul de l'offset time utilisent un offset time fixe. Cette méthode souffre d'un manque de robustesse en présence de trafic variable et nécessite une paramétrisation fine.

3.6.4 Priorité des trafics

Yoo et Qiao ont proposé dans [25] un mécanisme permettant d'isoler des classes de service par la détermination de l'offset time. Le mécanisme est expliqué plus en détail dans la section 3.8.1 et illustré par la figure 3.11. Il consiste à augmenter l'offset time des trafics prioritaires. L'inconvénient de la méthode réside dans le fait que le délai des trafics prioritaire est énormément allongé.

3.6.5 Utilisation de jetons aléatoires

Une idée proposée par [26] consiste à réguler l'envoi des rafales par un système de jetons. L'entête est envoyé dès qu'une rafale est prête. Celle-ci par contre devra attendre l'arrivée d'un jeton pour partir. Ainsi, l'arrivée des jetons, conjointement avec le mécanisme d'agrégation des rafales, détermine l'OT. Les auteurs proposent de rendre l'arrivée des jetons aléatoire.

3.6.6 Envoi agressif des paquets de contrôle

Liu et Ansari proposent dans [27] d'envoyer l'entête de signalisation avant que la rafale ne soit prête. Ils présentent pour cela une méthode permettant de prédire la taille de la rafale ainsi que la date d'envoi. L'entête est envoyé. Une fois que la rafale correspondante est prête, si la réservation de l'entête suffit pour la rafale, la rafale est envoyée. Dans ce cas, le délai d'accès de la rafale sera diminué. Sinon, un nouvel entête est envoyé conformément au protocole JET classique. La difficulté réside dans la fonction de prédiction de rafales qui, si elle n'est pas fidèle au comportement du trafic engendrera des réservations inutiles qui peuvent augmenter l'utilisation du réseau et dégrader les performances.

3.7 Initialisation des tables de routage

En mode OBS, l'intelligence est regroupée à l'accès. Une bonne configuration initiale permet donc de rendre les informations de l'accès plus pertinentes avant l'envoi des données. Parmi les décisions que le noeud d'accès devra prendre, le choix de l'OT est primordial et se base sur la taille du chemin à parcourir. Ensuite, plusieurs phénomènes liés à la congestion dans le réseau imposent des mécanismes plus élaborés. Une répartition de la charge sur le réseau permet de réduire les contentions.

Rouskas et Teng proposent dans [28] une modélisation amenant à du mono-routage (i.e., un seul chemin par connexion). Le choix du mono-routage est justifié par l'utilisation de la conversion de longueur d'onde comme unique moyen pour résoudre les contentions. Ainsi, sans déflexion, les rafales appartenant à une même requête emprunteront toutes le même chemin. L'objectif est de minimiser le taux de perte global du réseau. La fonction de coût de chaque lien est évaluée par une approximation par partie de la formule Erlang-B. Cette approximation permet de rendre la fonction objective linéaire. Le modèle de mono-routage peut être résolu avec une formulation en nombres entiers forçant les connexions à emprunter un seul chemin, ou encore en nombres réels offrant une solution de flots bifurqués. Avec la seconde approche, un mécanisme itératif permet ensuite de choisir un seul chemin par connexion. Cette alternative est intéressante car elle permet de réduire la complexité du problème dans le cas de gros réseaux. Dans ces travaux, la conversion de longueur d'onde est autorisée, ce qui justifie le choix absolu du mono-routage.

Huang, Heritage et Mukherjee proposent quant à eux ([29]) une fonction de coût dépendante de la taille du lien, de son utilisation et du nombre de chemins empruntant ce lien. La fonction objective consiste ensuite à minimiser la somme des coûts de chaque lien. L'évaluation des métriques sur une fenêtre de temps permet de rendre le mécanisme adaptatif.

3.8 Prise en charge de la qualité de service

La diversité des applications attendues pour la prochaine génération de l'Internet ainsi que la variété des profils de trafic font apparaître des classes de trafic qui, pour une bonne cohabitation, nécessitent la prise en charge de la qualité de service. La littérature propose plusieurs mécanismes que nous présentons dans cette section.

3.8.1 Utilisation de l'offset time

Yoo et Qiao ont montré dans [25] que le fait d'ajouter un *Extra Offset Time* (EOT) à une rafale permet d'augmenter sa priorité sans pour autant altérer les performances du trafic global sur le réseau.

L'illustration 3.11 permet de saisir en quoi un EOT (qui signifie le retardement de la rafale par rapport à son paquet de contrôle) permet d'isoler un flot parmi deux.

Supposons que P_1 et P_2 se soient vu allouer le même OT (scénario A). Dans ce cas, le premier paquet de contrôle (P_1) pourra réserver le lien et le second (P_2) sera perdu.

En ajoutant un EOT suffisamment grand à P_2 (il faut en fait que $OT_{P_2} > B_{P_1} + OT_{P_1}$ où B_{P_1} représente la taille de la rafale associée à P_1), on peut assurer qu'elle ne pourra être perdue à cause d'une rafale de priorité inférieure.

Si P_1 arrive avant P_2 (scénario B), il n'y aura pas contention et P_2 sera servi.

En fait, pour qu'une contention se produise (scénario C), il faut que P_2 arrive avant P_1 , auquel cas c'est P_1 qui sera perdu.

Dans tous les cas, l'EOT assure à P_2 la priorité sur P_1 . Cette technique peut être étendue à plus que deux classes de service ([24]).

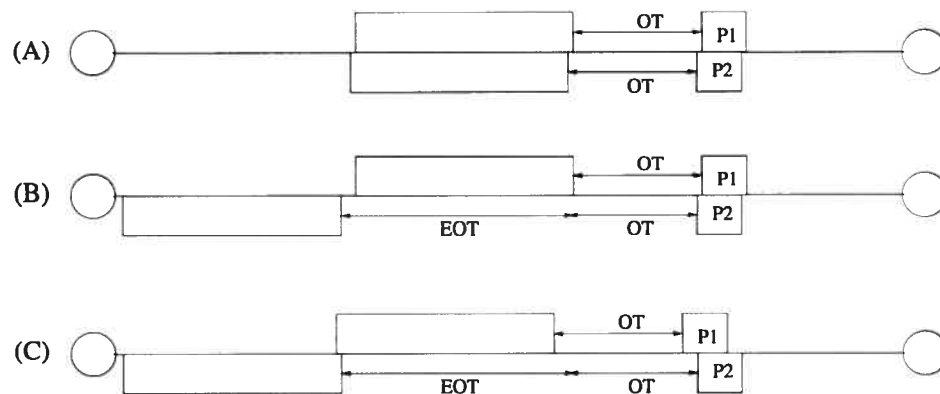


FIG. 3.11 – Illustration de l'isolation d'une classe de service par EOT

3.8.2 Utilisation de la segmentation

La segmentation consiste à diviser la rafale en contention. Dès lors, seule la partie en concurrence avec une autre rafale sera traitée comme une rafale en contention, l'autre partie pouvant utiliser le lien par défaut. La technique présentée dans [23] est élaborée conjointement au routage par déflexion pour combiner les deux techniques.

Ainsi, les auteurs définissent plusieurs stratégies :

- DP - "Drop Policy" : une des deux rafales est jetée.
- DDP - "Deflect and Drop Policy" : tente de défléchir une rafale, sinon la jette.
- SDP - "Segment and Drop Policy" : segmente une rafale puis jette la partie en contention.
- SDDP - "Segment and Deflect and Drop Policy" : segmente une rafale puis tente de défléchir la partie en contention. Sinon la jette.
- DSDP - "Deflect and Segment and Drop Policy" : défléchit une rafale ou tente de jeter la partie en contention.

Dès lors, le choix de la stratégie utilisée dépendra de la classe de service des rafales en contention.

3.8.3 Perte anticipée

En présence de plusieurs trafics, on peut délibérément choisir de jeter une rafale non prioritaire pour avantager les trafics prioritaires. Cette technique, proposée dans [30] permet d'assurer le service des classes prioritaires.

En notant d_{C_i} le nombre de rafales perdues pour la classe C_i et a_{C_i} le nombre de rafales arrivées sur un intervalle de temps F , les auteurs définissent la probabilité de perte de la classe i sur F par $p_{C_i} = d_{C_i}/a_{C_i}$. De plus, à chaque classe C_i est associé un taux de perte maximum $p_{C_i}^{\max}$ et un taux de perte minimum $p_{C_i}^{\min}$.

Deux stratégies sont alors proposées pour déterminer la perte anticipée de paquets :

EDT (Early Dropping by Threshold)

Tant qu'une classe i ne répond pas à ses exigences en termes de taux de perte, les rafales de la classe $i + 1$ (i.e., directement moins prioritaire) seront jetées.

$$p_{C_i}^{ED} = \begin{cases} 0 & p_{C_{i-1}} < p_{C_{i-1}}^{\max} \\ 1 & p_{C_{i-1}} \geq p_{C_{i-1}}^{\max} \end{cases} .$$

EDS (Early Dropping by Span)

Cette solution tend à anticiper le fait qu'une classe de trafic i dépasse $p_{C_i}^{\max}$. Une zone d'anticipation est définie par $P_{C_i}^{\min}$ et $P_{C_i}^{\max}$.

$$p_{C_i}^{ED} = \begin{cases} 0 & p_{C_{i-1}} < p_{C_{i-1}}^{\min} \\ \frac{p_{C_{i-1}} - p_{C_{i-1}}^{\min}}{\delta_{C_{i-1}}} & p_{C_{i-1}}^{\min} < p_{C_{i-1}} < p_{C_{i-1}}^{\max} \\ 1 & p_{C_{i-1}} \geq p_{C_{i-1}}^{\max} \end{cases} .$$

Chapitre 4

Accès au réseau OBS

En mode OBS, l'intelligence est regroupée à l'accès. Les décisions critiques pour le bon cheminement d'une rafale sont prises en fonction des informations disponibles au noeud d'accès et celles-ci seront irrévocables. En particulier, le calcul de l'OT, sans l'utilisation de fibre à retardement sera critique, de même que, sans utilisation de conversion de longueur d'onde, le choix de la longueur d'onde. Nous traitons ces deux questions en proposant tout d'abord quatre métriques de choix de longueurs d'ondes, puis un mécanisme adaptatif pour le calcul de l'OT.

4.1 Choix de la longueur d'onde

Sans conversion de longueur d'onde, le choix de la longueur d'onde sera effectué uniquement à l'accès et sera irrévocable.

Cet aspect est donc primordial pour le système et trois facteurs sont à considérer.

- La répartition de la charge sur les longueurs d'onde : sans convertisseurs de longueurs d'onde, le choix est définitif. Etant donné que le noeud d'accès ne dispose que d'informations limitées sur l'état du réseau, il semble judicieux d'essayer de

- répartir la charge sur les différentes longueurs d'onde.
- Le délai d'accès : lors du choix de la longueur d'onde, la rafale est agrégée mais les données sont toujours dans le domaine électrique. De ce fait, elles peuvent être retardées pour donner plus de flexibilité au mécanisme de choix et pour assurer qu'il n'y ait pas de contention sur le premier lien. Cela génère cependant un délai si le choix porte sur une longueur d'onde occupée.
 - La fragmentation de la capacité des longueurs d'onde (section 3.5.3).

4.1.1 Max-SV

La stratégie Max-SV consiste à choisir la longueur d'onde qui n'a pas été utilisée depuis le plus longtemps. Elle présente l'avantage de pouvoir être mise en oeuvre avec très peu de ressources. De plus, cette stratégie permet de répartir la charge sur les longueurs d'onde.

4.1.2 La longueur d'onde la moins utilisée sur le lien (LLWoL)

Cette stratégie consiste à utiliser la longueur d'onde la moins utilisée en termes de bande passante sur une fenêtre de temps coulissante. Elle demande plus de ressources que Max-SV mais assure la répartition de la charge de manière plus rigoureuse.

4.1.3 La longueur d'onde la moins utilisée par le noeud (LLWoLN)

Les paquets qui sont insérés au niveau d'un noeud d'accès vont créer de la concurrence pour les paquets en transit par ce noeud. De ce fait, il paraît judicieux de considérer la charge d'une longueur d'onde sur un noeud, c'est-à-dire la charge moyenne de cette longueur d'onde sur les liens. Cette métrique permet de laisser de la place sur la longueur d'onde afin de favoriser les chances de succès de la déflexion.

4.1.4 La longueur d'onde la moins utilisée par le noeud voisin (LL-WoNN)

On peut aussi considérer la métrique précédente sur le noeud suivant, afin d'offrir à la rafale plus de possibilités une fois convertie en signal lumineux.

4.2 Choix de l'offset time

L'offset time est un paramètre très important pour la conception de réseau OBS. Il doit être suffisamment grand pour assurer le traitement de l'entête et la réservation des ressources jusqu'à la destination. Ainsi, les pertes par IOT seront diminuées. D'un autre côté, l'OT est la plus grande partie du délai subi par la rafale, car celle-ci doit attendre que l'OT soit écoulé avant d'être envoyée. Il faut donc éviter de trop gros OT. De plus, de gros OT offrent aux rafales de circuler plus longtemps dans le réseau ce qui risque de surutiliser les ressources et d'augmenter la congestion ([31]).

Nous proposons de comparer trois stratégies de calcul d'OT.

4.2.1 Stratégie utilisant le chemin par défaut (S1)

Les noeuds d'accès connaissent un chemin par défaut pour chaque destination. Cette première stratégie consiste à déterminer l'EOT proportionnellement à la taille de ce chemin par défaut N_d .

$$EOT = N_d \times w$$

avec $w \geq 0$. Cette solution purement statique présente le seul avantage de la simplicité. Les performances dépendront uniquement de l'initialisation de w et la méthode souffre d'un manque certain de robustesse en présence de trafic dynamique.

4.2.2 Stratégie utilisant la taille moyenne des chemins (S2)

En ajoutant un mécanisme de retour d'information à la source, la stratégie précédente peut être adaptée pour améliorer ses performances. Comme $EOT \geq 0$, le chemin emprunté par une rafale est potentiellement plus long que celui par défaut. De plus, l'allongement des chemins empruntés précède les périodes de congestion. Ainsi, en définissant l'OT non plus comme une proportion du chemin par défaut, mais plutôt comme une proportion de la taille du chemin moyen, l'algorithme devient plus adaptatif, surtout avec un scénario dynamique.

En notant R_{sd} l'ensemble des chemins de S à D , ℓ_r la taille du chemin r et B_r^F le nombre de rafales ayant emprunté le chemin r sur une fenêtre F définie comme suit : la fenêtre F est définie sur un certain nombre d'intervalles N_F contigus de taille P . En d'autres termes, les informations sont conservées sur une durée $N_F \times P$. Dès qu'un intervalle de temps égal à P s'est écoulé, la fenêtre coulisse, c'est-à-dire que les $N_F - 1$ plus récents intervalles de la fenêtre deviennent les $N_F - 1$ plus anciens et l'intervalle courant est réinitialisé. Dans cette stratégie, une fenêtre de temps coulissante contient la somme des tailles des chemins empruntés et une autre contient le nombre de rafales reçues. La taille moyenne est calculée comme suit :

$$\bar{N}_{sd} = \frac{\sum_{r \in R_{sd}} B_r^F \ell_r}{\sum_{r \in R_{sd}} B_r^F}.$$

Dès lors, l'EOT peut être exprimé par la relation suivante :

$$EOT = \bar{N}_{sd} \times w.$$

4.2.3 Stratégie purement adaptative (S3)

Nous voulons rendre l'algorithme le moins dépendant possible de la paramétrisation. Il faut donc déterminer l'EOT uniquement à partir des performances et de l'état du réseau. Considérons un trafic T accédant au réseau par le noeud S et sortant du réseau

par D . L'OT est calculé sur la base du chemin par défaut auquel S ajoute un Extra Offset Time (EOT) mis à jour avec une période P . Lorsqu'une rafale de T est perdue, un message (NACK) est envoyé à S qui peut alors tenir à jour une fenêtre de temps coulissante (comme celle définie dans 4.2.2) contenant les taux de perte de T définis comme la probabilité de perte d'une rafale (indépendamment de la taille de celle-ci). Avec une structure similaire, D conserve la somme des tailles des chemins empruntés par les rafales arrivées à destination sur le même horizon et la somme des charges des noeuds traversés. Toutes les informations nécessaires à la destination sont récoltées par le paquet de contrôle.

A chaque mouvement de fenêtre, D envoie à S la somme des tailles des chemins utilisés pour T sur F (soit $\sum_{r \in R_{sd}} B_r^F l_r$) ainsi que la somme des charges des noeuds traversés et le nombre de rafales arrivées à destination : $\sum_{r \in R_{sd}} B_r \sum_{v_i \in r} LOAD_F(v_i)$, où $LOAD_F(v_i) = \frac{\sum \text{trafic entrant au noeud } v_i \text{ sur la durée } F}{deg(v_i) \times W \times C}$, avec W est le nombre de longueur d'onde pour chaque lien et C leur capacité.

Nous pouvons ainsi exprimer la charge d'un flot sd sur une fenêtre F par

$$LOAD_{sd}^F = \frac{\sum_{r \in R_{sd}} B_r (\sum_{v_i \in r} LOAD_F(v_i))}{\sum_{r \in R_{sd}} B_r l_r}.$$

Une première métrique liée à la charge des noeuds utilisés par le flot est définie par la relation

$$EOT_{load} = N_{sd}^{moy} \times LOAD_{sd}^F.$$

Ensuite, une seconde métrique permet d'adapter l'EOT au taux de perte.

$$EOT_{loss} = \tau_{IOT_{sd}}^F \times w_3.$$

où $\tau_{IOT_{sd}}^F = \sum_{v \in r: r \in R_{sd}} \tau_{IOT_{sd}}^F(v)$ est la somme des taux de perte causées par OT insuffisant pour le trafic entre S et D à chaque noeud v traversé sur la durée de F et $w_3 > 0$ est un poids représentant la sensibilité au taux de perte.

Les deux métriques sont reliées par une combinaison linéaire :

$$EOT_F = 2 \times (\gamma EOT_{load} + (1 - \gamma) EOT_{loss})$$

où γ permet de pondérer le poids de la métrique basée sur la charge par rapport au poids de la métrique sur la perte IOT.

Enfin, au temps $t \in F$, l'EOT sera calculé ainsi :

$$EOT_t = \beta EOT_F + (1 - \beta) EOT_{F-1}$$

où β est un coefficient de réminiscence et t appartient à F . Il indique à quel point le passé doit compter dans le choix de l'EOT, relativement à la valeur déterminée à partir de la fenêtre courante.

4.2.4 Choix de la date d'émission de la rafale

Etant donné qu'à l'accès, les rafales appartiennent encore au monde électrique, elles peuvent être retardées pour s'assurer qu'il n'y ait pas de contention sur le premier lien. Nous utiliserons donc un algorithme de seuil percé pour réguler le trafic en fonction de la disponibilité du lien d'entrée. En fait, la disponibilité du lien aura une influence sur le choix de la longueur d'onde.

Chapitre 5

Routage en mode OBS

5.1 Routage par déflexion

5.1.1 Algorithme de routage

L'algorithme de routage est exécuté pour chaque paquet de contrôle afin de déterminer le port de sortie en chacun des noeuds parcourus. Chaque noeud détient une table de routage et une procédure de déflexion. Les informations nécessaires, arguments de la procédure sont transportées par le paquet de contrôle. L'algorithme 1 décrit la procédure de routage.

5.1.2 Algorithme de déflexion

Si le chemin par défaut n'est pas libre pour la rafale, alors l'algorithme de déflexion est exécuté. Les routeurs optiques disposent de tables de routage de taille $(deg(v) \times N)$ où N est le nombre de noeuds et $deg(v)$ est le nombre de liens sortants du noeud v . Ainsi, pour chaque lien ℓ adjacent au noeud courant et chaque destination d , la table

Algorithme 1 Algorithme de routage

si \mathcal{R}_d est libre **alors**

$\mathcal{L}_b \leftarrow R_d$

sinon

si certains liens sont libres sur I et permettent d'atteindre d en moins de OTL noeuds **alors**

$\mathcal{L}_b \leftarrow \text{déflexion}(b)$

sinon

jeter(b)

finsi

finsi

contient la taille du chemin le plus court commençant par ℓ et menant à d ainsi qu'un indicateur de routage. La déflexion est par défaut basée sur la taille du chemin, avec une priorité inversement proportionnelle à celle-ci. L'indicateur de routage permet ensuite de départager les chemins de même taille. Nous proposons plusieurs métriques dans la suite du rapport.

5.1.3 Métriques

Le choix du chemin alternatif priorise la taille du chemin. Cependant, cet argument ne dépend que de la topologie du réseau et non de son état de congestion. Pour différencier les chemins de même taille, nous proposons plusieurs métriques.

- MAX-SV : l'indicateur MAX-SV a été introduit par Xu, Qiao, Li et Xu dans [21]. C'est à dire que, pour une longueur d'onde λ , parmi plusieurs liens, nous choisirons celui dont la longueur d'onde λ n'a pas été utilisée depuis le plus longtemps.
- Le lien le moins utilisé (LLL) : l'utilisation de la longueur d'onde λ du lien ℓ sur une fenêtre de temps F de durée d_F est définie par

$$U_{\ell,\lambda}^F = \frac{\text{temps réservé}}{d_F}.$$

Algorithme 2 Algorithme de déflexion

pour tout ℓ sortant de v **faire**

si $\text{longueur}(\mathcal{P}(\ell, d)) \leq OTL$ **alors**

si $\text{longueur}(\mathcal{P}(\ell, d)) < \text{longueur}(\mathcal{P}(\ell^*, d))$ **alors**

$\ell^* \leftarrow \ell$

fin

si $\text{longueur}(\mathcal{P}(\ell, d)) = \text{longueur}(\mathcal{P}(\ell^*, d))$ **alors**

si $\text{metrique}(\ell, d) < \text{metrique}(\mathcal{P}(\ell^*, d))$ **alors**

$\ell^* \leftarrow \ell$

fin

fin

fin

fin pour

retourne ℓ^*

- Le noeud le moins chargé (LLN) : la charge d'un noeud est définie par la relation suivante

$$\text{LOAD}^F(v) = \frac{\text{trafic reçu par } v \text{ sur } d_F}{\text{degres}(v) \times C_\lambda \times W},$$

où C_λ est la capacité d'une longueur d'onde et W le nombre de longueurs d'onde par lien. Nous choisirons donc le lien menant au noeud le moins chargé.

- Le noeud présentant la longueur d'onde la moins chargée (LLWN) : avec les mêmes notations que précédemment,

$$\text{LOAD}_\lambda^F(v) = \frac{\text{trafic reçu par } v \text{ sur } d_F \text{ sur } \lambda}{\text{degré}(v) \times C_\lambda}.$$

5.2 Initialisation des tables de routage

5.2.1 Motivations

Pour chaque flot, nous définissons un chemin par défaut. Ce chemin sert tout d'abord de base pour le calcul de l'*offset time*. Le chemin par défaut n'est plus consulté par le paquet de contrôle une fois qu'une déflexion a eu lieu. Une approche de type MPLS (label switching) réduit le temps de processus des paquets de contrôle et donc l'*offset time* et le délai bout en bout. Le *label switching* consiste à affecter des étiquettes (*label*) à chaque entête. Les noeuds utilisent ces étiquettes pour le routage ; pour chaque étiquette sur une interface d'entrée, ils associent une interface de sortie et un étiquette de sortie. Ainsi, les étiquettes permettent de définir un chemin choisi par la source.

Il est donc important de choisir l'ensemble des chemins par défaut de façon à minimiser le taux de contention.

5.2.2 Chemin le plus court

Une première approche consiste à choisir le chemin par défaut parmi les plus courts chemins. Parmi plusieurs chemins de même taille, nous choisirons le chemin qui présente la plus grande connectivité, définie comme suit :

$$\mathcal{R}_{sd} = \max_{r \in R_{sd}^{SR}} \sum_{v \in r} \text{deg}(v).$$

5.2.3 Modélisation du problème de flot

La modélisation de la topologie et du trafic sous forme de problème de flot permet de répartir les chemins par défaut afin de diminuer les contentions. Une flot est défini par un noeud source, un noeud destination et un débit. Nous proposons un modèle générique et le testons avec plusieurs fonctions objectives.

Notations

$C_\ell \in \mathbb{R}^+$	capacité d'une longueur d'onde du lien ℓ
W_ℓ	nombre de longueurs d'onde du lien ℓ
$A_f \in \mathbb{R}^+$	débit demandé par le flot f
$x_{f,\ell} \in [0, 1]$	part du flot f empruntant le lien ℓ
$u_\ell = \sum_f (x_{f,\ell} \times A_f) \in \mathbb{R}^+$	utilisation du lien ℓ
$u_\ell^{RES} = C_\ell \times W_\ell - u_\ell \in \mathbb{R}^+$	capacité résiduelle du lien ℓ
$u_n = \frac{\sum_{\ell \in \ell_n^-} u_\ell \times C_\ell \times W_\ell + \sum_{f \in f_n^-} A_f - \sum_{f \in f_n^+} A_f}{\text{card}(\ell_n^+)} \in \mathbb{R}^+$	utilisation du noeud n
ℓ_f^-, ℓ_f^+	liens dont la source (resp. destination) est la source (resp. dest) de f
ℓ_n^-, ℓ_n^+	liens entrant (resp. sortant) dans le noeud n
f_n^-, f_n^+	flots entrant (resp. sortant) dans le noeud n
$F_{*,d}$	ensemble des flots à destination de d
$F_{s,*}$	ensemble des flots en provenance de s
$\text{card}(E)$	cardinal de l'ensemble E

Objectif :

$$z^\ell = \alpha \max z_1^\ell + (1 - \alpha) \min z_2^\ell,$$

$$z^n = \alpha \max z_1^n + (1 - \alpha) \min z_2^n$$

où :

$$z_1^\ell = \sum_{\ell \in L} u_\ell,$$

$$z_2^\ell = \min_{\ell \in L} u_\ell^{RES},$$

$$z_1^n = \sum_{n \in N} u_n,$$

$$z_2^n = \min_{n \in N} -u_\ell^n.$$

En minimisant la fonction z_1^ℓ , la solution vise à minimiser l'utilisation globale des ressources, donc tend à utiliser les plus courts chemins. Cependant, cet objectif dépend uniquement de la topologie. De ce fait, il peut orienter les flots vers des chemins hautement partagés et créer des goulots d'étranglement. Une fonction d'équilibrage semblable mais plus complète (tenant compte notamment de la taille des liens) a été développée dans des travaux indépendants par Huang, Heritage et Mukherjee dans [29].

Maximiser la fonction z_2^ℓ permet d'éliminer les goulots d'étranglement en répartissant les flots sur les liens. La contrepartie est l'allongement des chemins et donc l'augmentation des délais.

Lorsque la fonction z_2^ℓ est utilisée, nous exécutons l'algorithme itératif 3 pour répartir le trafic sur l'ensemble du réseau et pas seulement au niveau du premier goulot d'étranglement. Il consiste à ajouter une contrainte permettant de forcer le nouveau goulot d'étranglement de la solution courante à sa valeur pour l'exécution suivante. Ainsi, à chaque exécution, un lien de plus est affecté comme goulot d'étranglement et à la fin de l'algorithme, tous les liens ont été vus comme des goulots d'étranglement et la répartition est totale.

Les fonctions z_1^n et z_2^n sont utilisées de la même manière que les fonctions précédentes mais en relation avec la charge des noeuds plutôt que la charge des liens. La fonction z^ℓ (resp. z^n) est une combinaison linéaire de z_1^ℓ (resp. z_1^n) et de z_2^ℓ (resp. z_2^n). Le poids α est utilisé pour augmenter le poids relatif de z_1^ℓ (resp. z_1^n) par rapport à z_2^ℓ (resp. z_2^n).

Contraintes

$$u_\ell \leq C_\ell \times W \quad , \ell \in L \quad (5.1)$$

$$\sum_{\ell \in \ell_n^+} x_{f,\ell} - \sum_{\ell \in \ell_n^-} x_{f,\ell} = \begin{cases} 0 & n \text{ est un noeud de transit} \\ -1 & f \in F_{*,d} \\ 1 & f \in F_{s,*} \end{cases} \quad , n \in N, f \in F. \quad (5.2)$$

Algorithme 3 Algorithme itératif pour l'utilisation de l'objectif z_2

pour tout lien faire

résoudre le problème de flot

$goulot = \ell$ tel que ($\ell \notin goulotS$) et $\forall t, t \notin goulotS, u_\ell^{RES} \leq u_t$

ajouter la contrainte $u_{goulot} = *u_{goulot}$ pour fixer l'utilisation du goulot lors des prochaines itérations.

ajouter $goulot$ dans $goulotS$

fin pour

L'équation (5.1) impose la contrainte de capacité. Notons que nous n'avons pas considéré les longueurs d'ondes et que nous nous assurons seulement du fait que les liens ne soient pas surutilisés. En effet, la granularité de l'OBS n'impose pas qu'un flot transite sur une longueur d'onde dédiée.

L'équation (5.2) représente la contrainte de conservation stipulant que dans chaque noeud et pour chaque flot, le trafic entrant est égale au trafic sortant, sauf pour les noeuds source (1) et destination (-1).

Bornes

$$0 \leq x_{f,\ell} \leq 1 \quad , f \in F, \ell \in L.$$

5.2.4 Utilisation de la solution du problème de flot

Le problème de flot nous offre pour chaque lien ℓ et chaque flot f la proportion de f qui passe par ℓ . Il reste désormais à construire les chemins correspondants.

Nous avons investigué deux approches.

Routage par la source

Pour chaque couple sd présentant au moins une requête, le problème de flot permet d'extraire plusieurs chemins utilisés. Nous choisissons le plus utilisé et l'imposons à l'envoi du paquet de contrôle. Une fois les chemins déterminés, ils peuvent être imposés aux paquets de contrôle soit en indiquant le chemin dans ceux-ci, soit en effectuant une distribution d'étiquettes (label) pour permettre une commutation de type MPLS. Le choix du mono-routage est lié à la procédure de calcul de l'OT. Les retours d'informations sont plus précis car plus nombreux et plus simples à mettre en place dans le cadre de mono-routage

$$r(s, d) = \max_{r \in R_{sd}} \min_{\ell \in r} \sum_{f \in F_{s, * \times F_{*, d}}} x_{f, \ell}.$$

Routage par la destination

Si la table de routage est initialisée en fonction de la destination uniquement, sa taille est réduite. Nous choisissons, à un noeud n et pour une destination d , le lien qui dirige le plus de trafic

$$\ell(n, d) = \max_{\ell \in \ell_n^+} \frac{\sum_{f \in F_{*, d}} x_{f, \ell} \times A_f}{\sum_{\ell, f \in F_{*, d}} x_{f, \ell} \times A_f},$$

où le numérateur représente l'utilisation du lien ℓ sortant de N par les flots à destination de D et le dénominateur représente l'utilisation de tous les liens sortant de N par des flots à destination de D .

Chapitre 6

Simulation et résultats

Nous simulons le réseau OBS sur un réseau d'échelle réelle avec des matrices réalistes. Dans cette section, nous testons toutes les propositions faites dans les sections 4 et 5 à savoir le choix de la longueur d'onde et du lien de déflexion, le calcul de l'offset time et l'initialisation des tables de routage. Nous utilisons uniquement la déflexion pour la résolution des contentions.

6.1 Description des expériences

6.1.1 Topologie du réseau

Nous utilisons le réseau EONET proposé par [3] pour faire des simulations en mode OBS-JET. La figure 6.1 représente le réseau coeur. Il est composé de 28 noeuds pour 41 arcs.

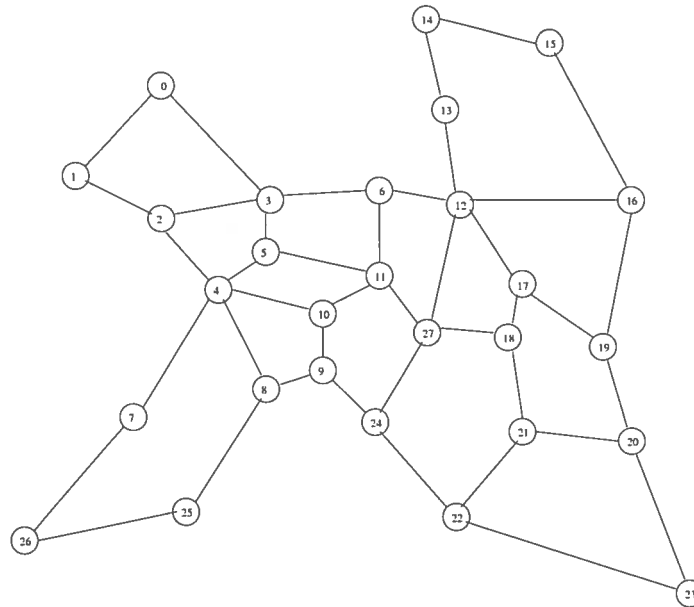


FIG. 6.1 – Le réseau EON

6.1.2 Les modèles de flots utilisés

Modèle uniforme

Nous avons simulé le réseau OBS-JET sous une matrice de requête uniforme.

$$(\forall i, j), T_{i,j} = \begin{cases} 0 & \text{si } i = j \\ C & \text{sinon} \end{cases} .$$

Matrice EONET_cost266

Pour des simulations plus réalistes, nous avons aussi utilisé la matrice de trafic proposée par [3]. Elle représente des évaluations de l'utilisation réelle du réseau en 2004. Cependant, elle représente un trafic symétrique pour chaque couple de noeuds et une charge totale de l'ordre de 4.3Tbps.

Trafic perturbé

Pour certaines simulations, nous utilisons un trafic présentant une perturbation. Les perturbations seront alors décrites dans la section correspondant à l'expérience.

6.1.3 Les caractéristiques observées

Les critères de qualité seront le taux de perte et le délai moyen bout en bout.

6.2 Expérimentation des stratégies de choix de longueurs d'onde

6.2.1 Description

Les fibres sont configurées à 50 longueurs d'ondes OC-192. Nous utilisons la matrice de trafic EONET_cost266 modifié par un coefficient multiplicateur (l'axe des abscisses) afin de faire varier la charge du réseau. Les quatre stratégies sont définies dans la section 4.1. La figure 6.2 représente le taux de perte obtenu après stabilisation du réseau pour plusieurs instances de charge, c'est à dire que le taux de perte reste constant et que les trafics ont atteint un équilibre.

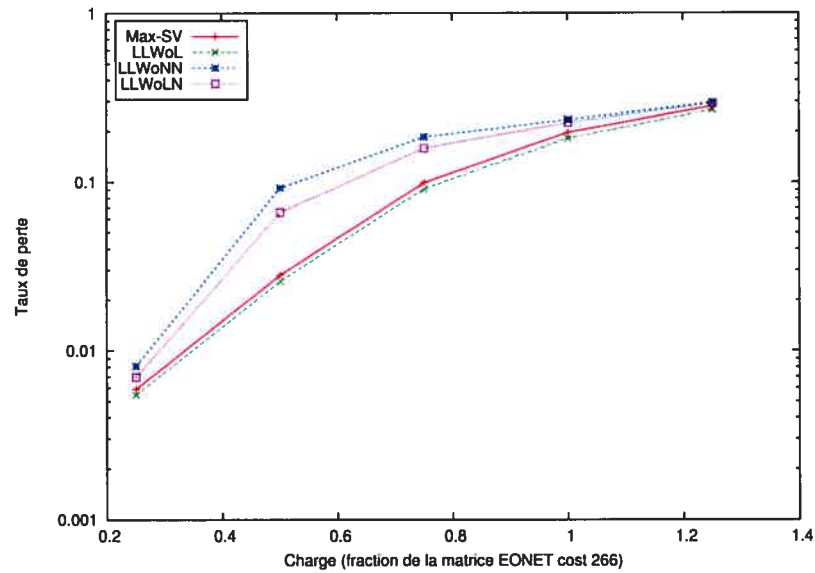


FIG. 6.2 – Taux de perte pour les stratégies de choix de longueur d’onde

6.2.2 Commentaires sur les stratégies de choix de longueurs d’onde

Les stratégies basées sur la charge d’une longueur d’onde sur un noeud, que ce soit sur le noeud courant LLWoLN ou sur le prochain noeud LLWoNN ne donnent pas de bons résultats comparativement à celles relatives au lien (LLWoL et Max-SV). De plus, aller chercher des statistiques sur le prochain noeud ne semble pas porter car LLWoNN est la moins bonne stratégie.

Les deux stratégies basées liens (Max-SV et LLWoL) se démarquent en termes de taux de perte. Si l’utilisation minimum de la longueur d’onde semble la meilleure métrique, la stratégie Max-SV est pratiquement équivalente et s’avère moins coûteuse, autant en termes de mémoire que de temps de calcul. Si ces aspects sont critiques, on choisira donc Max-SV. Sinon, on choisira l’utilisation minimum sur un horizon (LLWoL).

6.3 Expérimentation des stratégies de déflexion

6.3.1 Description

Nous effectuons ici une série de simulations basées sur la variation de la charge. Comme dans l'expérience précédente, nous utilisons 50 longueurs d'onde OC-192 par lien et la matrice EONET_266 modifiée par un coefficient (l'axe des abscisses). A l'accès, le choix de la longueur d'onde est effectué avec la métrique MAX-SV.

La figure 6.3 représente le taux de perte obtenu après stabilisation du taux de perte dans le réseau pour plusieurs instances de charge.

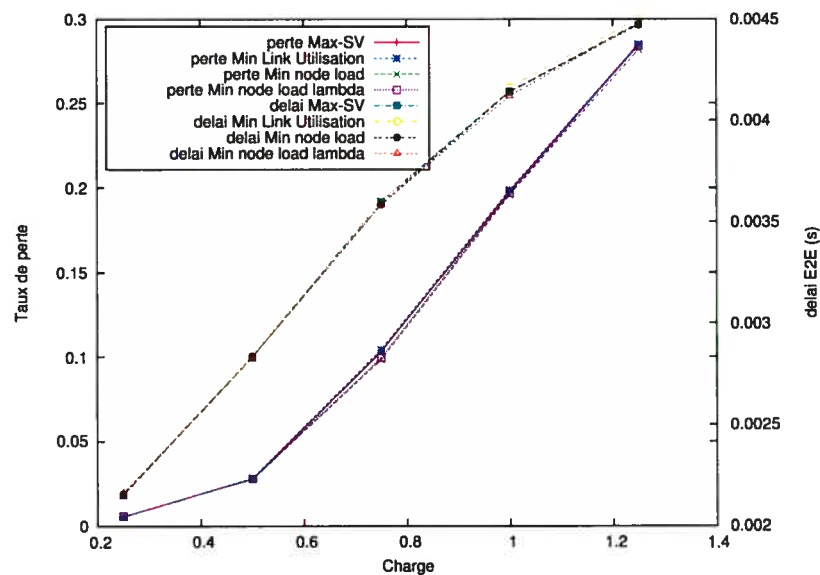


FIG. 6.3 – Taux de perte pour les stratégies de déflexion

6.3.2 Commentaires sur les stratégies de déflexion

Metnani a montré dans [31] que la déflexion priorisant le chemin le plus court est la meilleure stratégie par rapport à des métriques locales indépendantes de la taille

du chemin (comme les indicateurs que nous étudions). Notons que dans notre étude, nous avons modifié l'architecture du noeud d'accès. Metnani [31] utilise en effet un noeud supplémentaire pour chaque noeud d'accès qui est chargé de l'agrégation et de la régulation du trafic.

La modélisation explicite de l'OT rend cette architecture mal adaptée car l'accès en deux phases divise les connaissances et le choix de la date d'émission relativement à la valeur de l'OT n'est plus possible. Nous avons donc regroupé les fonctionnalités du noeud intermédiaire proposées par Metnani [31] dans le noeud d'accès. Avec cette architecture, aucun indicateur pour choisir parmi les chemins les plus courts ne se dégage, que ce soit en termes de taux de perte ou en termes de délai.

Une proposition faite par Yang et Rouskas dans [32] consiste à choisir un chemin plutôt qu'un lien. Ils proposent ainsi plusieurs métriques pour déterminer le coût d'un chemin et donc choisir le lien. Le mécanisme permet de tenir compte à la fois des métriques et de la taille des chemins et améliore les résultats. L'inconvénient de cette pratique réside dans la mise à jour des informations qui doivent être fréquentes et couvrir de longues distances.

6.4 Expérimentation des politiques de calcul de l'offset time

6.4.1 Influence des paramètres

Description

Nous étudions ici l'influence des paramètres γ et β de la stratégie adaptative présentée dans la section 4.2.3. Le réseau est dimensionné à 50 longueurs d'ondes OC-192. Pour le paramètre γ , nous utilisons une matrice de trafic uniforme avec $C = 423Mbps$. L'étude consiste à faire varier à la fois la charge du réseau et la valeur de γ . Les courbes 6.4 et 6.5 représentent respectivement le taux de perte et le délai et font apparaître des courbes de niveau pour plus de lisibilité.

Le paramètre β représente la réminiscence du passé et il intervient uniquement en présence de trafic variable. Nous utilisons la matrice EONET_cost266 avec un scénario consistant à augmenter progressivement le trafic (augmentation de 20% toutes les 10ms entre $t = 100$ et $t = 200$), puis, après une phase de trafic stable, à le faire décroître (diminution de 20% toutes les 10ms entre $t = 500$ et $t = 600$).

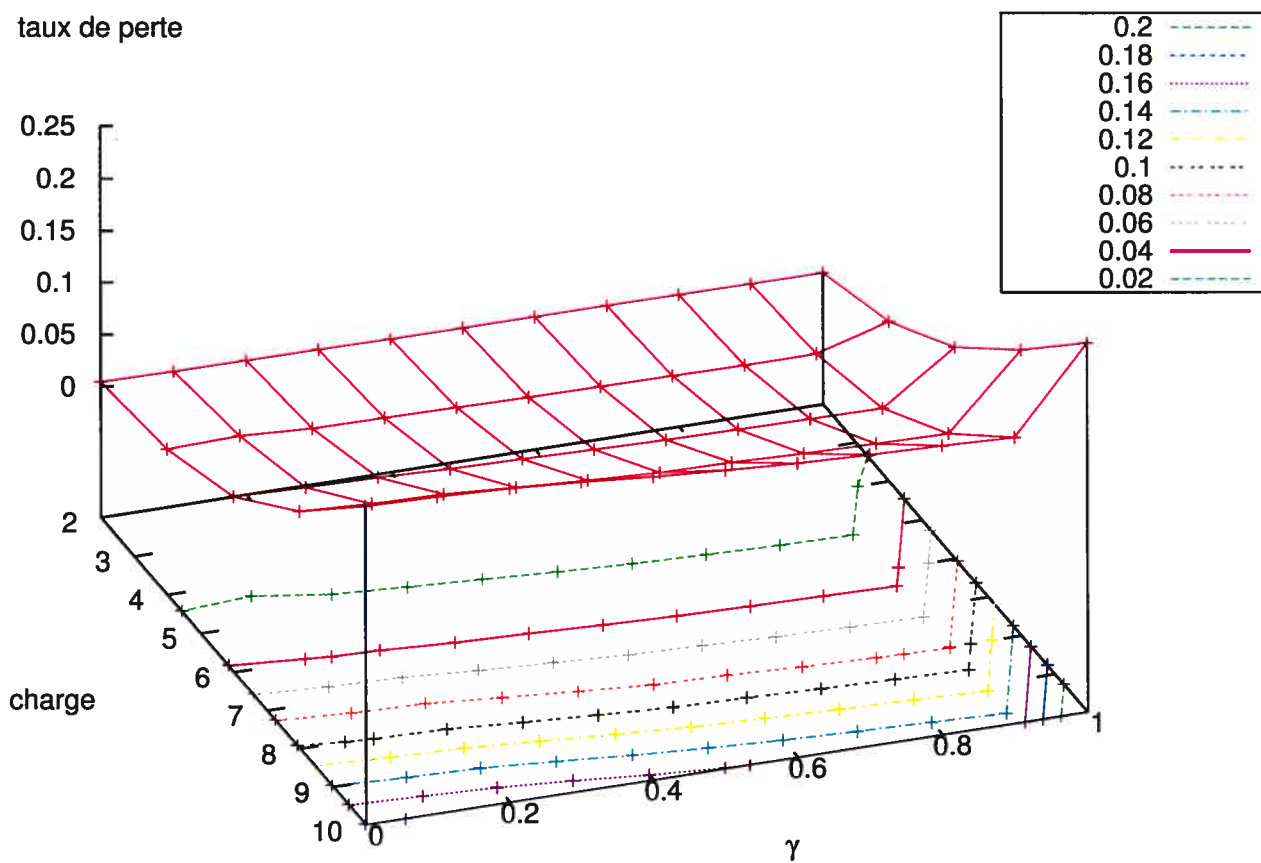


FIG. 6.4 – Influence du paramètre γ sur le taux de perte

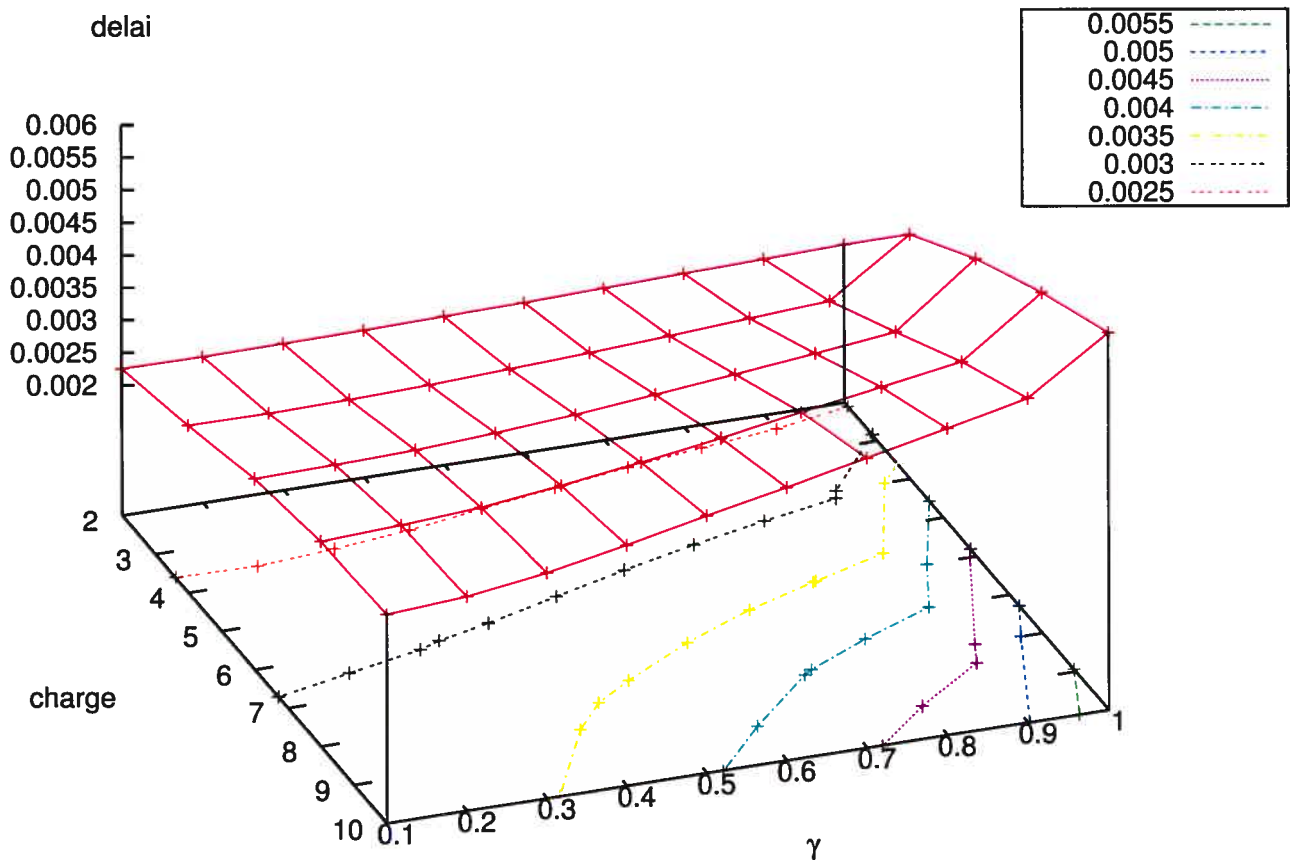


FIG. 6.5 – Influence du paramètre γ sur le délai

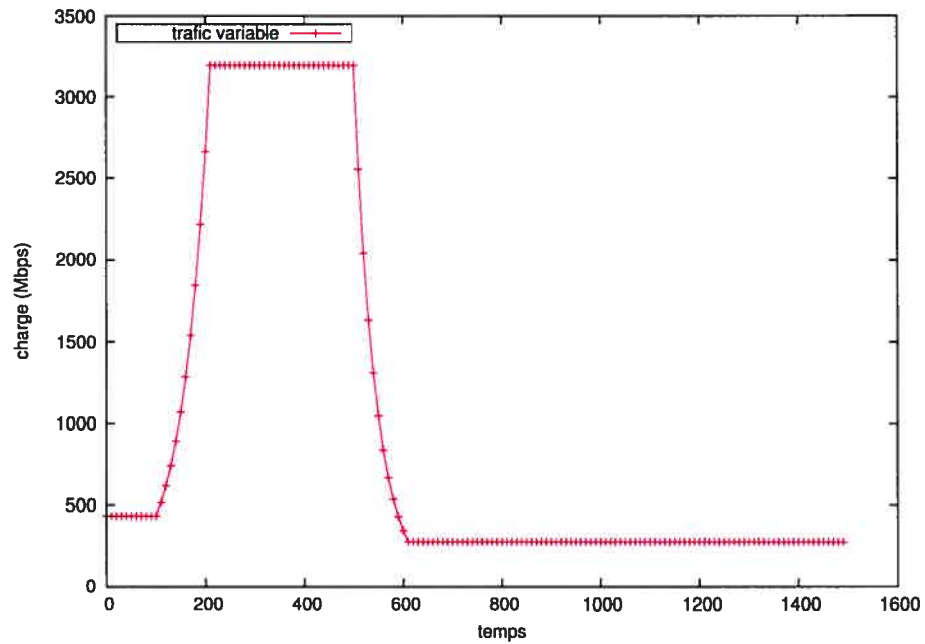


FIG. 6.6 – trafic variable utilisé pour l'étude du paramètre β

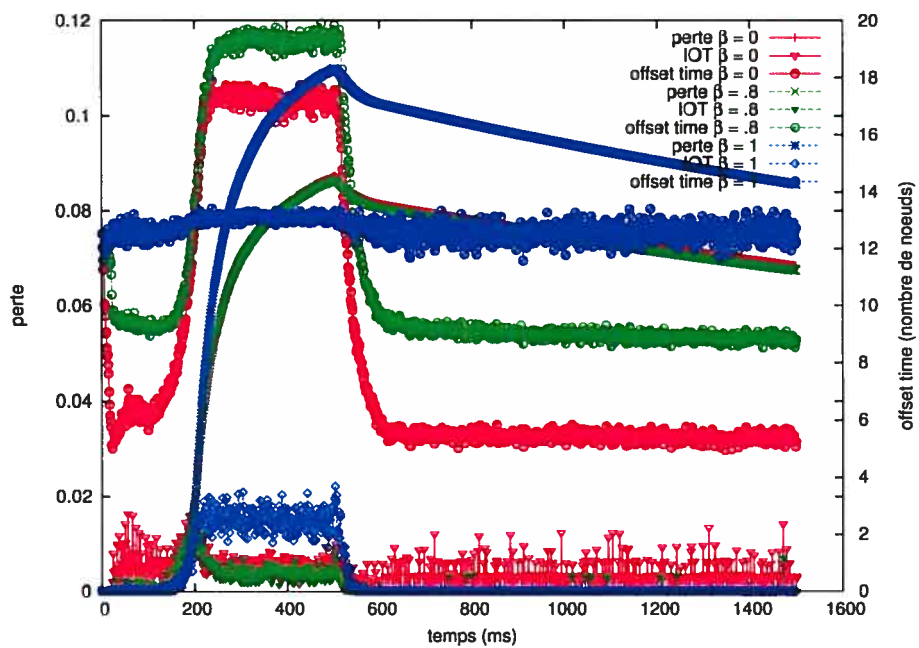


FIG. 6.7 – Influence du paramètre β

Analyse

En présence de trafic stable, le paramètre γ permet de jouer sur le compromis perte/délai. Si ne tenir compte que de l'indicateur de charge n'est pas pertinent ($\gamma = 1$) tant en termes de délai que de taux de perte, dès que les deux critères sont pris en compte (i.e. perte par IOT et charge), augmenter γ permet de diminuer les pertes en augmentant le délai. γ devra donc être choisi en fonction des caractéristiques des trafics mis en jeu.

Le coefficient de réminiscence β permet d'imposer une certaine rigidité au choix de l'OT :

- si $\beta = 1$, l'OT sera constant dans le temps et donc ne s'adaptera pas aux variations de trafic. La figure 6.7 permet de se rendre compte des conséquences : entre $t = 200ms$ et $t = 500ms$, l'OT est trop faible et le taux de perte augmente fortement. Ensuite, l'OT est trop grand et, si le taux de perte par IOT est annulé, du délai pourrait être économisé.
- si $\beta = 0$, l'OT varie rapidement sur une grande amplitude permettant de minimiser le taux de perte.
- si $\beta = .8$, l'OT varie rapidement moins rapidement que si $\beta = 1$ mais avec pratiquement autant d'amplitude sur l'expérience présentée.

Plus β est grand et moins le mécanisme aura de flexibilité pour s'adapter rapidement aux perturbations de trafic. Pour du trafic sensible au délai, on peut préférer affecter un grand β pour que son délai ne soit pas influencé par les variations de charge du réseau. Un trafic sensible au taux de perte quant à lui devra s'adapter à la charge du réseau le plus vite possible, préférant ainsi un faible β .

6.4.2 Expérience 1

Description

Nous avons expérimenté les trois stratégies de choix d'offset sur le réseau EONET. Les liens sont dotés de 32 longueurs d'ondes OC-192. Nous utilisons un trafic uniforme avec $C = 423Mbps$ dont l'intensité varie entre chaque simulation. La courbe 6.8 illustre le taux de perte et le délai bout en bout pour une charge variant de 1 à 11. Pour la stratégie adaptative, $\gamma = 0.2$, $\beta = 0.5$ et $\omega = 300$.

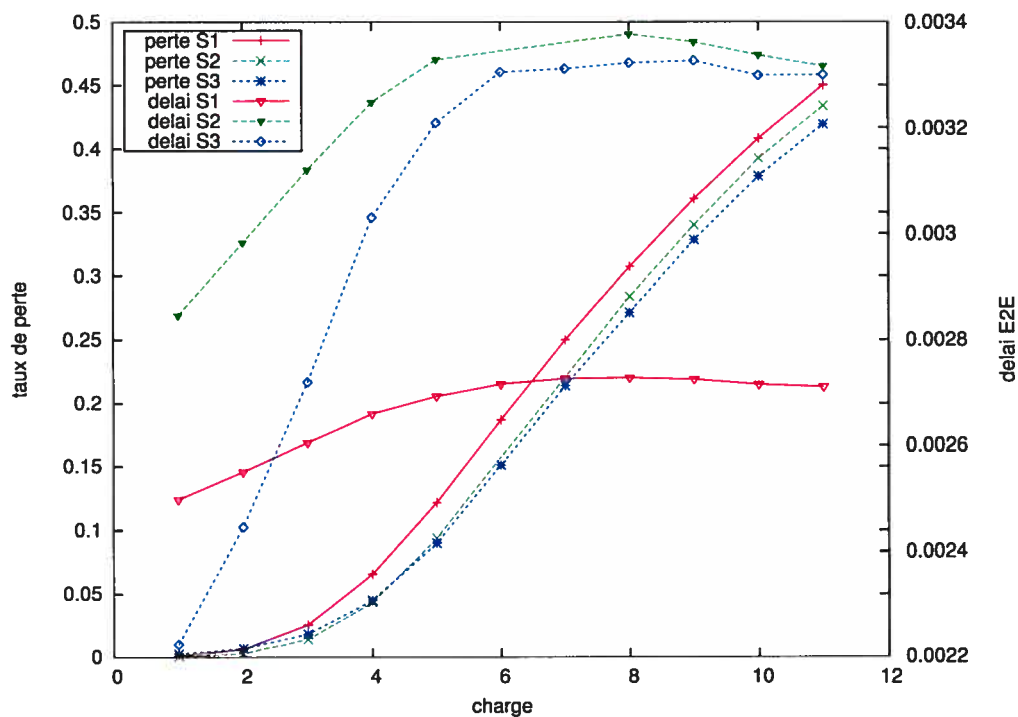


FIG. 6.8 – Les stratégies d'offset time

Analyse

De manière générale, les stratégies de calcul de l'OT à partir du chemin par défaut et du chemin moyen n'offrent pas une forte amplitude dans les valeurs calculées. Ainsi, à faible charge, la stratégie adaptative peut réduire le délai bien au dessous des autres stratégies pour un taux de perte pratiquement équivalent. Ensuite, lorsque la charge augmente, toutes les stratégies augmentent l'OT et la stratégie adaptative donne le meilleur taux de perte avec un délai inférieur à celui obtenu avec le chemin moyen. La stratégie du chemin par défaut assure un délai pratiquement constant et ne s'adapte donc pas à l'augmentation du trafic. Il en résulte le pire taux de perte pour le chemin par défaut.

6.4.3 Expérience 2

Description

Nous désirons maintenant mettre en évidence l'atout principal de la stratégie adaptative : le fait de s'adapter à une variation de trafic. Nous utilisons toujours le réseau EONET, mais avec la matrice `cost_266`. Le trafic est initialisé à la moitié de la matrice originale puis, au bout de $60ms$, il est multiplié par 2. Enfin, au temps $200ms$, le trafic est multiplié par 0.75. Les résultats sont illustrés par la figure 6.9.

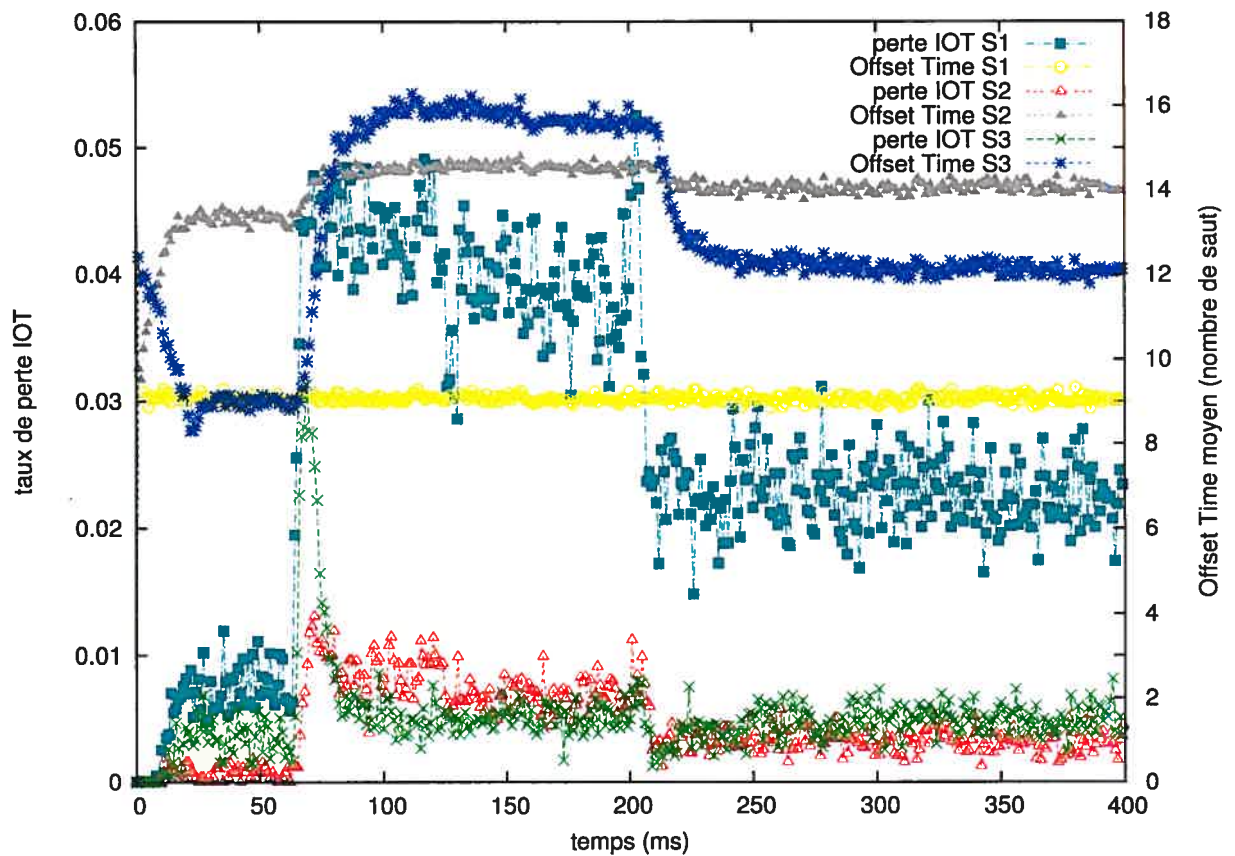


FIG. 6.9 – Evolution du taux de perte par IOT et de l'offset time aux perturbations de trafic

Analyse

Sur la première partie de la figure 6.9 ($t \leq 60ms$), la stratégie S3 garantit un taux de perte par IOT de l'ordre de 0.5% avec le plus petit OT. Le taux de perte par IOT obtenu avec la stratégie S2 est légèrement inférieur mais avec un OT bien plus élevé. Lorsque le trafic double, la stratégie S3 augmente brusquement l'OT pour garantir le même taux de perte dû à l'IOT qu'avant la perturbation (0.5%). S2 quant à elle ne dispose pas de la même amplitude de variation, mais peut tout de même allonger l'OT pour atteindre un taux de perte de l'ordre de 1%.

Enfin, lorsque le trafic diminue ($t \geq 200ms$), S2 et S3 adoptent un OT intermédiaire comparativement aux deux premières parties de l'expérience. Là encore, S3 parvient à conserver un taux de perte IOT constant ce qui lui permet de réduire l'OT. Le taux de perte IOT de S2 est comparable pour un OT plus important.

6.4.4 Conclusion sur les politiques d'offset time

La stratégie S1 offre un offset time constant quelque soit la charge et ne peut donc pas s'adapter aux perturbations. La stratégie S2 peut moduler la valeur de l'offset time avec les variations du trafic mais dans une amplitude restreinte qui ne lui permet donc pas de couvrir de grosses perturbations. La stratégie S3 peut s'adapter très rapidement à de grosses perturbations pour garder un taux de perte par IOT pratiquement constant.

6.5 Initialisation des tables de routage

Maintenant, nous allons évaluer la pertinence de notre modèle pour l'initialisation des tables de routage.

6.5.1 Description

Nous confrontons le taux de perte et le délai pour les cinq stratégies d'initialisation des tables de routage. Les simulations utilisent le trafic EONET_cost266 pour un réseau dimensionné avec 50 longueurs d'ondes OC-192 par lien.

Les figures 6.10 et 6.11 comparent les différentes stratégies d'initialisation respectivement pour un routage par destination et par chemin. La figure 6.12 quant à elle permet de confronter le routage par source et par destination pour les meilleurs critères.

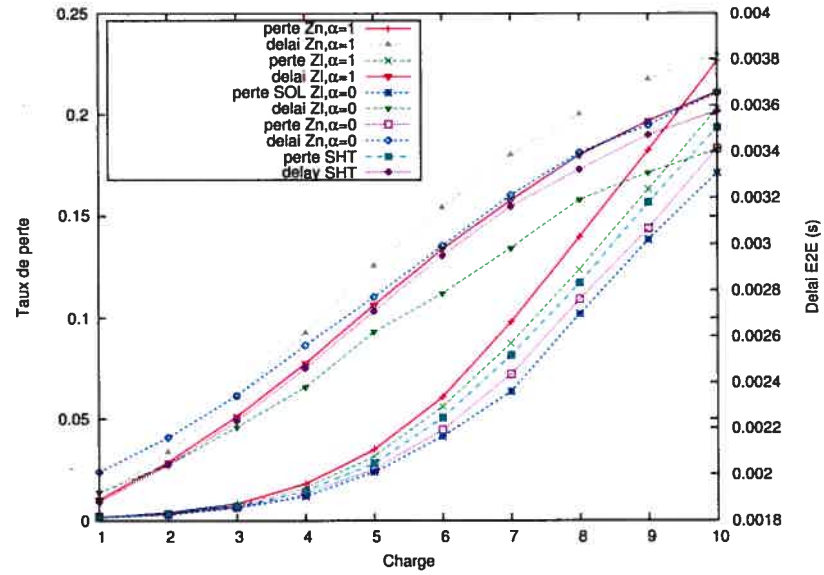


FIG. 6.10 – Délai et taux de perte avec un routage par la destination

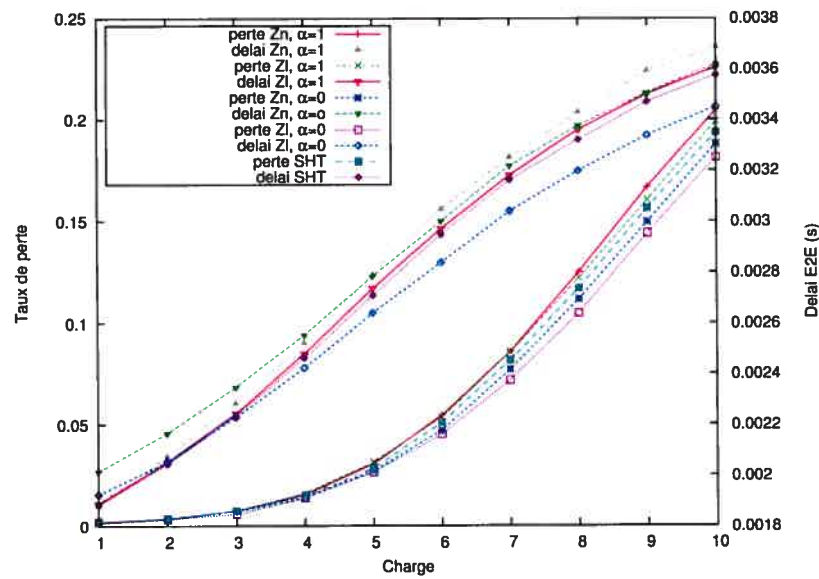


FIG. 6.11 – Délai et taux de perte avec un routage par la source

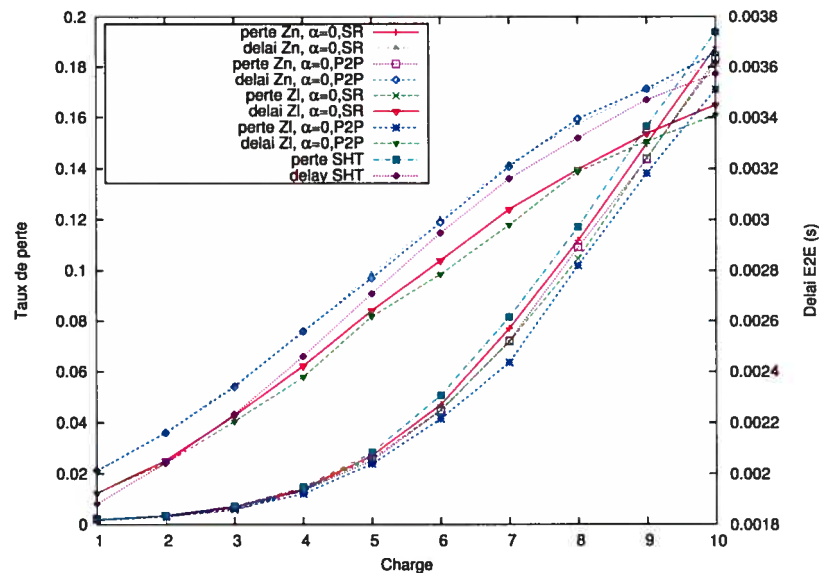


FIG. 6.12 – Comparaison du routage par la source et du routage par la destination

6.5.2 Analyse des résultats

Les stratégies consistant à étaler le trafic avec un objectif de type max min sont les meilleures en termes de compromis délai/perte. En considérant la métrique de lien, les performances sont les meilleures en termes de perte et de délai. Les stratégies minimisant la somme d'une métrique donnent de moins bons résultats que l'initialisation par les chemins les plus courts.

Si l'on essaie de comparer le routage par la source et le routage par la destination, la courbe 6.12 permet de donner clairement l'avantage au routage par la destination. Cela provient du fait que nous utilisons uniquement la déflexion pour résoudre les contentions. Dans un tel contexte, lorsqu'une déflexion apparaît, la rafale défléchie ne bénéficie plus de l'équilibrage de charge et utilise une métrique de routage basée sur la taille du chemin qui n'est pas la plus efficace.

On peut aisément concevoir que l'utilisation de conversion de longueur d'onde avantagerait grandement le routage par la source dans la mesure où la conversion permet

de suivre plus longtemps le chemin par défaut et donc de bénéficier de l'équilibrage de charge.

Chapitre 7

Conclusion

A l'heure où certains pensent à la fibre optique résidentielle, la conception des réseaux optiques reste très dépendante de l'évolution et des progrès de la couche physique. Le trafic résidentiel présentera des exigences de plus en plus fortes en termes de qualité de service et des caractéristiques de plus en plus variables étant donné les prochaines applications à prévoir. En effet, des trafics temps réel (télévision) devront cohabiter avec des trafics très irréguliers (téléphonie) et des applications plus classiques (web, transfert de fichiers).

La commutation de circuits (OCS), si elle présente l'avantage d'un contrôle plus puissant sur les performances, ne fait pas preuve d'une assez bonne flexibilité dans le cadre de trafics fortement variables ([33]). Pour améliorer les performances de OCS, il faudra compter sur la couche physique et améliorer les noeuds du coeur, ce qui implique des frais importants.

La commutation de paquets, en revanche, présente une excellente granularité et serait très bien adaptée à des trafics variables. Néanmoins, le matériel requis pour la conception pratique de l'OPS n'est pas encore disponible.

L'OBS a donc un potentiel crédible du fait qu'il nécessite un matériel n'appelant pas à autant d'investissements et d'innovations.

Cependant, le paradigme en tant que tel n'est pas mûr. Beaucoup de problèmes n'ont pas encore été traités. Nous en avons adressé quatre, à savoir le choix de la longueur d'onde et du lien de déflexion, le calcul de l'offset time et l'initialisation des tables de routage, plus ou moins populaires, et avons proposé des solutions tant à l'accès qu'à l'intérieur du réseau.

Mais quelle place les industriels et les fournisseurs sont-ils prêts à laisser à l'OBS ? Les paradigmes étant par essence très différents, ils s'adressent certainement à des profils de trafic et à des infrastructures différentes. Une étude comparative s'impose donc pour pouvoir identifier les scénarii favorables à un paradigme plutôt qu'à un autre. En effet, si l'adoption d'un paradigme alternatif à l'OCS peut effrayer par les chamboulements d'infrastructures sous-jacents, une cohabitation serait quant à elle tout à fait envisageable.

Bibliographie

- [1] E. Haselton. A PCM Frame Switching Concept Leading to Burst Switching Network Architecture. In *IEEE communication magazine*, volume 21, pages 13–19, June 1983.
- [2] C. Qiao and M. Yoo. Optical Burst Switching (OBS) - A New Paradigm for an Optical Internet. *Journal of High Speed Networks*, 8(1) :69–84, January 1999.
- [3] A. Betker, C. Gerlach, R. Hulsermann, M. Jager, M. Barry, S. Bodamer, J. Spath, C. Gauger, and M. Kohn. Reference Transport Network Scenarios. Technical report, MultiTeraNet Project, 2004.
- [4] T. Coutelen, H. Elbiaze, and B. Jaumard. An Efficient Adaptive Offset Mechanism to Reduce Burst Losses in OBS Networks, à paraitre. In *Globecom*, 2005.
- [5] T. Coutelen, H. Elbiaze, B. Jaumard, and A. Metnani. Impact of Network Topology and Traffic Load Balancing on Deflection Routing Efficiency. In *IASTED*, pages 132–137, 2005.
- [6] T. Coutelen, H. Elbiaze, B. Jaumard, and A. Metnani. Measurement-Based Alternative Routing Strategies in Optical Burst-Switched Networks. In *ICTON*, pages 224–227, 2005.
- [7] R. Ramaswami and K.N. Sivarajan. *Optical Networks, A Practical Perspective (Second Edition)*. Elsevier, 1998.
- [8] J.P. Jue and V.M. Vokkarane. *Optical Burst Switched Network*. Springer, 2005.
- [9] X. Ma and G.S. Kuo. Optical switching technology comparison : optical MEMS vs. other technologies. In *Communications Magazine, IEEE*, volume 41, pages S16,S23, November 2003.
- [10] F. Masetti, D. Zriny, D. Verchère, J. Blanton, T. Kim, J. Talley, D. Chiaroni, A. Jourdan, J.-C. Jacquinet, C. Coeurjolly, P. Poignant, M. Renaud, G. Eilenberger, S. Bunse, W. Latenshleager, J. Wolde, and U. Bilgac. Design and Implementation of a Multi-Terabit Optical Burst/Package Router prototype. In *Optical Fiber Communication Conference and Exhibit*, pages FD1–1,FD1–3, March 2002.
- [11] J. Moy. RFC-2328 : OSPF Version 2, April 1998.

- [12] B. Jaumard, C. Meyer, and B. Thiongane. ILP Formulations for the RWA Problem for Symmetrical Systems. To appear in Handbook of Telecommunication.
- [13] A. Jarray and B. Jaumard. Exact ILP Solution for the Grooming Problem in WDM Ring Network. In *IEEE ICC*, April 2005.
- [14] <http://www.cisco.com/en/US/products/hw/optical/ps2006/ps2010/>.
- [15] G. Bernstein, D. Caviglia, and R. Rabbat. Operating Virtual Concatenation (VCAT) and the Link Capacity Adjustment Scheme (LCAS) with Generalized Multi-Protocol Label Switching (GMPLS), July 2005.
- [16] X. Cao, J. Li, Y. Chen, and C. Qiao. TCP/IP Packets Assembly over Optical Burst Switching Network. In *IEEE Globecom*, volume 3, pages 2808,2812, November 2002.
- [17] A. Agustí-Torra, G.V. Bochmann, and C. Cervelló-Pastor. Intermediate Node Initiated (INI) signaling : A Hybrid Reservation Technique for Optical Burst-Switched Networks. In *2nd IFIP Intern. Conf. on Wireless and Optical Communications Networks (WOCN)*, March 2005.
- [18] J.Y. Wei and R.I. McFarland. Just-in-time Signaling for WDM Optical Burst Switching Networks. *Journal of Lightwave Technology*, 18(12) :2019–2037, December 2000.
- [19] J.P. Jue R. Karanam, V. M. Vokkarane. Intermediate Node Initiated (INI) signaling : A Hybrid Reservation Technique for Optical Burst-Switched Networks. In *Optical Fiber Communication Conference (OFC)*, March 2003.
- [20] J.S. Turner. Terabit Burst Switching. *Journal of High Speed Networks*, 8(1) :3–16, January 1999.
- [21] J. Xu, C. Qiao, J. Li, and G. Xu. Efficient Channel Scheduling Algorithms in Optical Burst Switched Networks. In *INFOCOM*, volume 3, pages 2268–2278, 2003.
- [22] Hunter, Cornwell, Gilfedder, Franzen, and Andonovic. SLOB : a Switch with Large Optical Buffers for. Packet Switching. In *IEEE/OSA Journal of Lightwave Technology*, volume 16, pages 1725–1736, October 1998.
- [23] V. Vokkarane and J. Jue. Burst Segmentation : an Approach for Reducing Packet Loss in Optical Burst Switched Networks. *Optical Networks Magazine*, 4(6) :81–89, Nov./Dec. 2003.
- [24] M. Yoo, C. Qiao, and S. Dixit. Optical Burst Switching for Service Differentiation in the Next-Generation Optical Internet. In *IEEE Communication Magazine*, pages 98–104, February 2001.
- [25] M. Yoo and C. Qiao. Supporting Multiple Classes of Services in IP over WDM Networks. In *GLOBECOM*, volume 1b, pages 1023–1027, 1999.
- [26] H. Chaskar, S. Verma, and R. Ravikanth. Robust Transport of IP Traffic over WDM using Optical Burst Switching. In *Nokia internal rep.*, January 2001.

- [27] J. Liu and N. Ansari. Aggressive Resource Reservation for OBS Systems. In *Communications, IEE Proceedings*, pages 233–238, August 2003.
- [28] J. Teng and G.N. Rouskas. Routing Path Optimization in Optical Burst Switched Networks. In *Proceedings of the 9th Conference on Optical Network Design and Modeling (ONDM 2005)*, pages 1–10, Milan, Italy, February 2005.
- [29] Y. Huang, J. P. Heritage, and B. Mukherjee. Dynamic Routing with Preplanned Congestion Avoidance for Survivable Optical Burst-Switched (OBS) Networks. In *Proc., OFC, OWC6*, March 2005.
- [30] Z. Zhang and B. Ramamurthy. Dynamic traffic grooming algorithms for reconfigurable SONET over WDM networks. *IEEE Journal on Selected Areas in Telecommunications*, 21(7) :1165–1172, September 2003.
- [31] A. Metnani. Routage par déflexion dans les réseaux optiques. Master's thesis, Université de Montréal, 2004.
- [32] G.N. Rouskas L. Yang. Adaptive Path Selection in Optical Burst Switched Networks . *soumis*.
- [33] T. Coutelen, H. Elbiaze, and B. Jaumard. Performance Comparison of OCS and OBS Switching Paradigms. In *ICTON*, pages 212–215, 2005.