



HAL
open science

Réseaux optiques en mode paquet pour les connexions internes à un centre de données

Barbara Pilar Dumas Feris

► **To cite this version:**

Barbara Pilar Dumas Feris. Réseaux optiques en mode paquet pour les connexions internes à un centre de données. Réseaux et télécommunications [cs.NI]. Ecole nationale supérieure Mines-Télécom Atlantique, 2017. Français. NNT : 2017IMTA0057 . tel-01878508

HAL Id: tel-01878508

<https://theses.hal.science/tel-01878508>

Submitted on 21 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

**UNIVERSITE
BRETAGNE
LOIRE**

THÈSE / IMT Atlantique

sous le sceau de l'Université Bretagne Loire

pour obtenir le grade de

DOCTEUR D'IMT Atlantique

Spécialité : Télécommunications

École Doctorale Mathématiques et STIC

Présentée par

Barbara Pilar Dumas Feris

Préparée dans le département Optique

**Réseaux optiques en mode
paquet pour les connexions
internes à un centre de données**

Thèse soutenue le 22 décembre 2017

devant le jury composé de :

Angélique Rissons

Professeur, Isae - Toulouse / présidente

Catherine Lepers

Professeur, Télécom SudParis / rapporteur

Philippe Chanclou

Ingénieur R&D, Orange Labs - Lannion / rapporteur

Philippe Gravey

Directeur d'études, IMT Atlantique / examinateur

Ammar Sharaiha

Professeur, Ecole nationale d'ingénieurs de Brest / directeur de thèse

Jelena Pesic

Ingénieur R&D, Nokia BELLS - Nozay / invitée

Marie-Laure Moulinard

Ingénieur d'études, IMT Atlantique / invitée

Michel Morvan

Ingénieur d'études, IMT Atlantique / invité

Pascal Morel

Maître de conférences, Ecole nationale d'ingénieurs de Brest / invité

Remerciements

Je tiens à remercier tous les membres du département optique de Télécom Bretagne ainsi que les membres du Labsticc à l'ENIB. Je remercie de même les membres du projet EPOC.

Tout d'abord j'exprime ma gratitude envers le Professeur Pierre Pellat-Finet, qui travaillait à l'ambassade française à Santiago, et qui m'a mise en contact avec le Directeur du département optique de Télécom Bretagne, le Professeur Jean-Louis De Bougrenet De la Tocnaye, pour faire un Master et continuer mon stage chez eux.

Je remercie tout particulièrement mes encadrants : les professeurs Philippe Gravey, Marie-Laure Moulinard, Michel Morvan, Pascal Morel et mon Directeur de thèse : Professeur Amar Sharaiha. Chacun de mes encadrants a su me guider sous des perspectives différentes, ce qui a enrichi mon apprentissage et le résultat final de ma thèse. Un grand merci à vous tous.

Je remercie également les amis du département optique : Jennifer Romer, Maina Sinou, Julien Le Meur, Kevin Raimondeau, Stephen Le Gall, Jean-Baptiste Lamour, Yoran-Eli Pigeon, Di Zhang, Charbel Nassour, Lida Sadeghioon et Ion Popescu.

Je remercie les enseignants, les doctorants et les amis de l'ENIB.

Je remercie aussi d'autres enseignants et amis de Télécom Bretagne pour les bons moments passés ensemble : Jelena Pesic, Ahmed Triki, Aimée Johansen, Cristina Arenas Buendía, José Lugo, Manuel Aranaz Padrón, Sara Pinedo, Sergei Poltaretski, Elvis Puche, Adriana Antonelli, Ana Pucheu, Isabella Rinco Tilieri et Monique Calmon.

Je remercie les enseignants de la Universidad Federico Santa María de Valparaíso, en particulier aux professeurs : Ricardo Olivares Véliz, Walter Grote Hahn, le regretté Héctor Carrasco Espinosa, Sergio Olavarría Simonsen, Wolfgang Freund Grunewaldt et Reinaldo Vallejos Campos. Je tiens à remercier aux professeurs : Alejandra Beghelli Zapata et Ariel Leiva López de la Universidad Adolfo Ibáñez et de la Pontificia Universidad Católica de Valparaíso.

Je remercie les amis du Chili : Sabina Torres Figueroa, Karenn Johnson Riquelme, Varinia Quiñones Olivares, Katherine Aliste Rosemberg, Catalina Tobar Bustos, Rafael Pedemonte Lavis et Fernando Toro Yany.

Je remercie la famille Schuster de m'avoir accueillie : Romain, Dominique, Frédéric, Marc, Nastasia et les grand-parents.

Je remercie ma famille : Margarita Feris Astorga (mère), Eduardo Dumas Feris (frère) et Leonor Astorga de Feris (grand-mère).

VERSIÓN EN ESPAÑOL

Agradezco a todos los miembros del departamento de óptica de Télécom Bretagne y los miembros de Labsticc de ENIB. Agradezco también a los miembros del proyecto EPOC (proyecto que financió el doc).

Para empezar agradezco al profesor Pierre Pellat-Finet, que en 2011 trabajaba en la embajada de Francia en Santiago y me contactó con el director del departamento de óptica de Télécom Bretagne, profesor Jean-Louis De Bougrenet De la Tocnaye, para hacer un segundo magíster y realizar una práctica de investigación con ellos.

Agradezco sinceramente a mis profesores guía : Philippe Gravey, Marie-Laure Moulinard, Michel Morvan, Pascal Morel y al director de tesis : profesor Ammar Sharaiha. Cada uno de ellos me supo guiar desde perspectivas diferentes, lo que enriqueció mi aprendizaje y el resultado final de la tesis de doctorado. Muchísimas gracias a todos ellos.

Agradezco a los amigos del departamento de óptica de Télécom Bretagne : Jennifer Romer, Maina Sinou, Julien Le Meur, Kevin Raimondeau, Stephen Le Gall, Jean-Baptiste Lamour, Yoran-Eli Pigeon, Di Zhang, Charbel Nassour, Lida Sadeghioon e Ion Popescu.

Agradezco a profesores, colegas y amigos de ENIB.

Agradezco también a otros profesores y amigos de Télécom Bretagne por haberme acompañado en esta etapa : Jelena Pesic, Ahmed Triki, Aimée Johansen, Cristina Arenas Buendía, José Lugo, Manuel Aranaz Padrón, Sara Pinedo, Sergei Poltaretski, Elvis Puche, Adriana Antonelli, Ana Pucheu, Isabella Rinco Tilieri y Monique Calmon.

Agradezco profundamente a los profesores de la Universidad Federico Santa María de Valparaíso, en particular a los profesores que jugaron un rol fundamental en mi formación como ingeniera e investigadora : Ricardo Olivares Véliz, Walter Grote Hahn, Héctor Carrasco Espinosa (Q.E.P.D), Sergio Olavarría Simonsen, Wolfgang Freund Grunewaldt y Reinaldo Vallejos Campos. Agradezco también a los profesores : Alejandra Beghelli Zapata y Ariel Leiva López de la Universidad Adolfo Ibáñez y de la Pontificia Universidad Católica de Valparaíso, con quienes aprendí y compartí varios momentos.

Agradezco a los amigos de Chile por sus consejos y buenos deseos : Sabina Torres Figueroa, Karenn Johnson Riquelme, Varinia Quiñones Olivares, Katherine Aliste Rosemberg, Catalina Tobar Bustos, Rafael Pedemonte Lavis y Fernando Toro Yany.

Agradezco a la familia Schuster por haberme acogido : Romain, Dominique, Frédéric, Marc, Nastasia y a los abuelos.

Agradezco a mi familia (los más importantes) por apoyarme a la distancia : Margarita Feris Astorga (mamá), Eduardo Dumas Feris (hermano) y Leonor Astorga de Feris (abuelita).

Réseaux optiques en mode paquet pour les connexions internes à un centre de données

RÉSUMÉ

La consommation d'énergie des centres de données est un enjeu majeur. Leurs communications internes représentent près du quart de cette consommation. Les technologies de commutation optique ont en principe une meilleure efficacité énergétique que les solutions actuelles. Ce travail porte sur les réseaux optiques en mode paquet pour des centres de données de petite et moyenne taille. Il s'est déroulé dans le cadre du projet EPOC (*Energy Proportional and Opportunistic Computing*) qui s'intéresse à la réduction de la consommation d'énergie d'un centre de données alimenté partiellement par des sources renouvelables. Une hypothèse clé est l'absence d'un réseau de stockage dédié réduisant ainsi la consommation des interconnexions. Par contre, afin de pouvoir éteindre certains serveurs selon la charge de travail et l'énergie disponible, le débit doit être de l'ordre de 100 Gbit/s. Après un état de l'art des réseaux optiques pour centre de données nous choisissons une solution reposant sur une infrastructure entièrement passive et des émetteurs rapidement accordables en longueur d'onde, proposée récemment dans la littérature (POPI).

Nous étudions les limitations dues aux composants (pertes d'insertion, plage d'accord et espacement des canaux). Nous proposons une extension (E-POPI) qui permet d'augmenter le nombre de serveurs connectés en utilisant plusieurs plages de longueurs d'onde. Pour les centres de données de plus grande taille nous proposons un réseau à deux étages (intra- et inter-racks) opérant respectivement dans les bandes C et L, POPI+. La connexion entre étages se fait via une passerelle transparente dans un sens et opaque dans l'autre. Différentes solutions de contrôle des deux étages sont détaillées.

Un des éléments essentiels de faisabilité de ces architectures est de concilier la montée en débit avec les pertes du réseau passif d'interconnexion. Les techniques cohérentes des transmissions longue distance ne sont pas actuellement envisageables pour un centre de données. Nous avons donc étudié les formats PAM 4 et 8, par simulation avec différents débits (jusqu'à 112 Gbit/s et récepteurs (PIN, APD et SOA-PIN) et aussi, expérimentalement, à 12 et 18 Gbit/s. Nous avons développé une méthode de compensation des distorsions générées par les différents composants qui procure un compromis entre précision de correction et temps de calcul.

Ces résultats nous permettent de déterminer les pertes d'insertion tolérables. Nous les combinons avec les limitations liées à la plage d'accord des émetteurs et à l'encombrement spectral des canaux occupant des fenêtres multiples de 12,5 GHz pour dimensionner les différentes architectures. Les réseaux POPI, E-POPI et POPI+ permettent respectivement la connexion de 48, 99 et 2352 entités à 112 Gbit/s. Nos évaluations tiennent compte d'une possible dispersion des caractéristiques des principaux composants.

Mots clés – Réseaux optiques, centre de données, commutation de paquets optiques, formats de modulation d'amplitude PAM 4 et PAM 8.

Packet-Optical Networks for Intra-Data-Center Connections

ABSTRACT

Data-center energy consumption is nowadays a major issue. Intra-data-center networking accounts almost for a quarter of the data-center total power consumption. Optical switching technologies could provide higher power efficiency than current solutions based on electrical-packet switching. This work focuses on optical-packet-switched networks for small- and medium-size data centers. It takes part of the EPOC (Energy-Proportional and Opportunistic Computing) project, which main interest consists on reducing the overall power consumption of a data center partially powered by renewable sources. A key assumption is that our data center does not rely on a dedicated storage network, in order to reduce the consumption of those interconnections. In addition, with the aim of being able to turn off some servers according to the workload and the available energy, the bit rate must be close to 100 Gbit/s. We have chosen, after studying the state of the art of data-center interconnects, a purely passive network architecture based on fast-wavelength-tunable transmitters under the name of POPI.

We study POPI's limitations due to its components (insertion loss, tuning range and channel spacing). We then propose an extension called E-POPI that allows to increase the number of connected servers by using several transmission bands. For larger data centers, we propose POPI+, a two-stage infrastructure for intra- and inter-rack communications operating in the C and L bands, respectively. The connection between both stages is done via a transparent gateway in one direction and an opaque one in the other. We discuss different control solutions for both stages.

The feasibility of these architectures depends on, among other factors, dealing with bit-rate increasing and power losses of a passive interconnect. Coherent long-distance-transmission techniques are not currently suited to data centers. We therefore studied PAM 4 and 8 modulation formats with direct detection. On the one hand, by simulation, with different bit rates (up to 112 Gbit/s and receivers (PIN, APD and SOA-PIN) and, on the other hand, experimentally, at 12 and 18 Gbit/s. We have developed a method for compensating the distortions generated by the different network components. Our method takes into account a good tradeoff between correction accuracy and computation time.

Simulation results allow us to determine the amount of insertion loss that may be supported. We then combine these results with the limitations of transmitters-tuning range and channel spacing using multiple of 12.5 GHz slots for dimensioning the proposed architectures. POPI, E-POPI and POPI+ interconnects allow the connection of 48, 99 and 2352 entities, respectively, at 112 Gbit/s. Our assessments take into account a potential dispersion of the characteristics of the main architecture components.

Keywords – Optical network, data center, optical-packet switching, pulse-amplitude modulation PAM 4 and PAM 8.

TABLE DES MATIÈRES

LISTE DES FIGURES	vii
LISTE DES TABLEAUX	xiii
1 INTRODUCTION GÉNÉRALE	1
2 ETAT DE L'ART	5
2.1 Introduction	5
2.2 Technologies optiques	5
2.2.1 Transmission sur fibre optique	5
2.2.2 Réseau et commutation optiques	18
2.2.3 Autres technologies optiques	20
2.2.4 Formats de modulation pour DC	21
2.3 Architectures d'interconnexion intra-DC	23
2.3.1 Centre de données actuels	23
2.3.2 Architectures d'interconnexion optiques	27
2.3.3 Consommation d'énergie des différentes architectures d'interconnexion	42
2.4 Description du projet EPOC	43
2.5 Conclusion	45
3 ÉTUDE DE L'ARCHITECTURE D'INTERCONNEXION POPI ET DE SON EXTENSION	47
3.1 Introduction	47
3.2 Analyse de POPI	47
3.2.1 Evaluation de la performance	47
3.2.2 Bilan de liaison	49
3.2.3 Variantes de POPI	52
3.2.4 Limite de capacité	56
3.3 POPI étendu : E-POPI	57
3.3.1 Options d'E-POPI à deux bandes	58
3.3.2 Options d'E-POPI à b bandes	64

3.4	Conclusion	65
4	ARCHITECTURE D'INTERCONNEXION À DEUX ÉTAGES : POPI+	67
4.1	Introduction	67
4.2	Description de l'architecture POPI+	67
4.2.1	Variante de POPI+ : Transfert-POPI	75
4.3	Bilan de liaison et dimensionnement de POPI+	80
4.4	Conclusion	83
5	APPLICATION DU FORMAT PAM AUX RÉSEAUX OPTIQUES ÉTUDIÉS	85
5.1	Introduction	85
5.2	Manipulations expérimentales	85
5.2.1	Banc de mesure	85
5.2.2	Résultats	93
5.3	Simulations sous VPItransmissionMaker 9.7	96
5.3.1	Schémas de simulation	96
5.3.2	Pré-compensation de la distorsion	101
5.3.3	Résultats de simulation	108
5.3.4	Capacité maximale de POPI, E-POPI et POPI+	113
5.4	Conclusion	121
6	CONCLUSIONS ET PERSPECTIVES	125
6.1	Conclusions	125
6.2	Perspectives	127
	BIBLIOGRAPHIE	129
	A GLOSSAIRE	141
	B DIMENSIONNEMENT DES RÉSEAUX POPI ET E-POPI	145
	C PUBLICATIONS	155

LISTE DES FIGURES

2.1	Schéma de principe des sources lasers accordables DBR [10].	7
2.2	(a) Structure SSG-DBR ; (b) profil de réflectivité [13].	7
2.3	(a) Structure SG-DBR ; (b) profil de réflectivité [19].	8
2.4	Schéma de principe du modulateur Mach Zehnder.	9
2.5	Taux d'extinction de l'EAM [33].	9
2.6	Atténuation en fonction de la longueur d'onde [35].	11
2.7	Système de transmission WDM à N canaux. MUX : multiplexeur, DEMUX : démultiplexeur.	17
2.8	Consommation d'énergie des formats de modulation NRZ, PAM, CAP et OFDM, à un débit binaire constant de 28 Gbit/s [59].	22
2.9	Schéma de modulation PAM : (a) émission (Tx) ; (b) réception (Rx).	22
2.10	Plusieurs cas de modulation PAM. Les files représentent le niveau de modula- tion M : PAM 2 (a, b, c) ; PAM 4 (d, e, f) ; et, PAM 8 (g, h, i). Les colonnes représentent les constellations (a, d, g) ; le spectre optique (b, e, h) ; et, le diagramme de l'œil (c, f, i).	24
2.11	Architecture d'interconnexion EPS classique.	25
2.12	Evolution du trafic dans les DC depuis l'année 1990, trafic exprimé en zéta-octet par an [70].	25
2.13	Tendances de distribution du trafic dans le DC, (a) publiées en 2013 [71], (b) publiée en 2016 [72].	26
2.14	Architecture d'interconnexion TWIN.	28
2.15	Architecture d'interconnexion inter-racks basée sur OCS classique : (a) schéma d'interconnexion, commutateurs OCS représentés en vert ; (b) connexions entre les couches d'agrégation et ToR ; (c) et (d) fonctionnement de l'OCS en utilisant des MEMS [94].	30
2.16	Architecture d'interconnexion Proteus.	31
2.17	Architecture d'interconnexion inter-baies HELIOS [77].	32
2.18	Architecture d'interconnexion inter-baies DOS [73]. LE : <i>Label Extractor</i> , TWC : <i>Tunable-Wavelength Converter</i> , O/E-E/O : <i>Optical to Electrical/ Electrical to Optical</i>	32

2.19	Architecture d'interconnexion inter-baies Petabit [73]. <i>IMs : Input Modules, CMs : Central Modules, OMs : Output Modules, TWC : Tunable-Wavelength Converter, AWGR : Arrayed-Waveguide-Grating Routing.</i>	33
2.20	Architecture d'interconnexion inter-baies IRIS [80] : (a) commutateur de longueur d'onde (<i>WS : Wavelength Switch</i>); (b) détecteur de l'entête (<i>HD : Header Detector</i>); (c) module de retard (<i>TB : Time Buffer</i>).	34
2.21	Architecture d'interconnexion <i>Broadcast and Select</i> [81].	35
2.22	(a) Architecture d'interconnexion basée sur OSS; (b) nœud OSS [82].	35
2.23	Architecture d'interconnexion POITR [83] : (a) 1 Rx; (b) 2 Rx. <i>OI : Optical Interface, WTT : Wavelength-Tunable Transmitter</i>	36
2.24	Architecture d'interconnexion EODCN [85] : (a) structure intra-rack; (b) structure inter-racks; (c) structure de contrôle. <i>ONI : Optical-Network Interface. BVT : Bandwidth-Variable Transceiver. TCAM : Ternary-Content-Addressable Memory.</i>	38
2.25	Architecture d'interconnexion POPI : (a) transmission des données; (b) transmission des messages de contrôle.	40
2.26	Réseau POPI : (a) interfaces de données et de contrôle; (b) exemple de connexion entre les serveurs 5 et 2, Tx_d et Tx_c représentent les émetteurs de données et de contrôle; Rx_d et Rx_c , les récepteurs de données et de contrôle, respectivement.	41
2.27	Consommation d'énergie des architectures d'interconnexion pour DC.	42
3.1	Performance de POPI en termes de : (a) latence moyenne; (b) débit moyen par serveur en fonction de la charge ρ_w [86].	49
3.2	Comparaison du bilan de liaison entre POPI et POITR.	51
3.3	POPI option B : (a) interfaces de données et de contrôle. (b) exemple de connexion entre les serveurs 5 et 2, Tx_d et Tx_c représentent les émetteurs de données et de contrôle; Rx_d et Rx_c , les récepteurs de données et de contrôle, respectivement.	53
3.4	POPI option C : (a) interfaces de données et de contrôle. (b) exemple de connexion entre les serveurs 5 et 2, Tx_d et Tx_c représentent les émetteurs de données et de contrôle; Rx_d et Rx_c , les récepteurs de données et de contrôle, respectivement.	54
3.5	POPI option D : (a) interfaces de données et de contrôle. (b) exemple de connexion entre les serveurs 5 et 2, Tx_d et Tx_c représentent les émetteurs de données et de contrôle; Rx_d et Rx_c , les récepteurs de données et de contrôle, respectivement.	55
3.6	Comparaison du bilan de liaison entre les options A, B, C et D de POPI.	56
3.7	Débit de chaque canal en fonction du nombre de serveurs dans le DC, dans le cas où la bande totale est limitée par la source accordable.	57

3.8	E-POPI $1-2-1-1$: (a) schéma d'interconnexion d'E-POPI $1-2-1-1$, $1, 2, \dots, n$: longueurs d'onde relatives aux serveurs $1, 2, \dots, n$, respectivement, r et g : longueurs d'onde relatives aux messages de <i>report</i> et de <i>grant</i> , respectivement ; (b) structure en émission et en réception.	59
3.9	E-POPI $1-1-2-1$: (a) schéma d'interconnexion, $1, 2, \dots, n$: longueurs d'onde relatives aux serveurs $1, 2, \dots, n$, respectivement, r et g : longueurs d'onde relatives aux messages de <i>report</i> et de <i>grant</i> , respectivement ; (b) structures en émission et en réception.	61
3.10	E-POPI $1-1-1-2$: (a) schéma d'interconnexion, $1, 2, \dots, n$: longueurs d'onde relatives aux serveurs $1, 2, \dots, n$, respectivement, r et g : longueurs d'onde relatives aux messages de <i>report</i> et de <i>grant</i> , respectivement ; (b) structure en émission et en réception.	63
4.1	Achitecture d'interconnexion TWIN basique basée sur un coupleur $N : 1$ et un démultiplexeur (DEMUX) $1 : N$. $1, 2, \dots, N$: longueurs d'ondes relatives aux entités (racks ou serveurs) $1, 2, \dots, N$, respectivement.	68
4.2	Achitecture d'interconnexion TWIN adaptée aux connexions inter-racks. Structure basée sur un coupleur $N + 1 : 1$, un amplificateur optique (OA) et deux étages de WSS $1 : N$. $1, 2, \dots, N, G$: longueurs d'ondes relatives aux racks $1, 2, \dots, N$ et au <i>gateway</i> du DC, respectivement.	70
4.3	POPI adaptée pour les connexions intra-racks du réseau à deux étages POPI+. (a) Schéma d'interconnexion. Connexions intra-racks en bande de transmission optique C (bleue) ; et inter-racks en bande de transmission L (rouge). Par simplicité nous avons omis les récepteurs de contrôle (toujours présents dans l'architecture). (b) Structure des serveurs en émission et en réception. (c) Structure de la passerelle du rack en émission et en réception.	72
4.4	Schéma d'interconnexion complet POPI+. Serveurs source (colonne de gauche) et serveurs de destination (colonne de droite). C : Coupleur ; $Buf.$: <i>Buffer</i>	73
4.5	Schéma d'interconnexion du réseau de contrôle de POPI+.	74
4.6	POPI adaptée pour les connexions intra-racks du réseau à deux étages T-POPI. (a) Schéma d'interconnexion. Connexions intra-racks en bande de transmission optique L (rouge) ; et inter-racks également en bande de transmission L (rouge). Par simplicité nous avons omis les récepteurs de contrôle (toujours présents dans l'architecture). (b) Structure des serveurs en émission et en réception. (c) Structure de la passerelle du rack en émission et en réception.	77
4.7	Structure des contrôleurs locaux en émission et en réception.	78
4.8	Variante de T-POPI : (a) schéma d'interconnexion intra-rack ; (b) structure en réception de la passerelle du rack.	79
4.9	Capacité maximale du réseau POPI+ : (a) puissance d'émission minimale pour le réseau POPI intra-rack ; (b) puissance minimale en entrée de la structure inter-racks.	82

5.1	Schéma bloc du banc de mesure d'un système PAM en détection directe.	86
5.2	Génération du signal PAM sous VPItransmissionMaker TM 9.7.	86
5.3	Signaux PAM 4 (a) et PAM 8 (b) bipolaires.	87
5.4	Caractérisation du MZM : puissance optique en fonction de la tension électrique appliquée.	88
5.5	Schéma bloc du banc de mesure d'un système PAM en détection directe avec pré-amplification optique (SOA).	89
5.6	Configuration de l'application KEYSIGHT TECHNOLOGIES sur l'OMA pour détecter les formats PAM 4 et 8.	90
5.7	Constellation de QAM 16 pour le calcul de l'EVM : (a) constellation idéale ou de référence ; (b) constellation simulée avec du bruit gaussien (EVM=15%) [132].	90
5.8	L'EVM classique contre l'EVM calculée par l'OMA.	91
5.9	Fonctions de densité de probabilité de chaque niveau pour le format de modulation PAM 4.	92
5.10	Résultats en termes d'EVM expérimental et par simulation, ainsi que de SER simulé, en fonction de la puissance optique reçue, pour le format de modulation PAM 4 à des débits symboles de 2, 3, 4 et 6 GBaud.	94
5.11	Résultats d'EVM expérimentale en fonction de la puissance optique reçue, pour les formats de modulation PAM 4 et 8, à une longueur d'onde fixe de 1550,12 nm, pour des débits symboles de 2, 3, 4 et 6 GBaud.	95
5.12	Résultats d'EVM expérimentale en fonction de la puissance optique reçue, pour les formats de modulation PAM 4 et 8, pour des longueurs d'onde de 1550,12, 1520,25 et 1570,42 nm, pour un débit symbole fixe de 6 GBaud.	95
5.13	Résultats de puissance optique reçue au niveau de l'OSA et d'EVM expérimentale en fonction de la longueur d'onde, pour le format de modulation PAM 4 à 6 GBaud, sans et avec pré-amplification optique à l'aide d'un SOA.	96
5.14	Schéma de simulation des réseaux POPI, E-POPI et POPI+ intra-rack : transmission optique de 5 canaux WDM avec le format de modulation PAM (4 ou 8).	98
5.15	Schéma de simulation modifié : insertion d'un ou deux EDFA sur la ligne de transmission optique pour (a) et (b), respectivement. Ces schémas sont utilisés pour simuler le réseau POPI+ inter-racks.	99
5.16	Fonction de transfert du module <i>ModulatorEA_Measured</i> de VPI : puissance optique en fonction de la tension électrique appliquée (figure de VPItransmissionMaker TM 9.7 modifiée afin de visualiser les paramètres <i>Bias</i> et <i>Drive</i> , ainsi que le signal PAM 4 en entrée du modulateur).	100
5.17	Schéma de simulation avec différents types de récepteurs optique : (a) photodiode PIN ; (b) photodiode APD ; (c) SOA + photodiode PIN.	102

- 5.18 Diagramme de l'œil de PAM. Les lignes représentent le niveau de modulation M : PAM 4 ((a), (b), (c) et (d)) et PAM 8 ((e), (f), (g) et (h)). Les figures sur les colonnes impaires ((a), (c), (e) et (g)) et paires ((b), (d), (f) et (h)) représentent le signal PAM électrique et le signal PAM optique à l'entrée et à la sortie du modulateur optique, respectivement. Les figures sur les 1^{ère} et 2^{ème} colonnes ((a), (b), (e) et (f)) représentent des signaux PAM sans méthode de pré-distorsion; sur les 3^{ème} et 4^{ème} colonnes ((c), (d), (g) et (h)), avec la méthode de pré-distorsion décrite ci-dessus. Paramètres utilisés : $Bias = -0,4/-0,3$ V, $Drive = -2/-1,95$ V et $\Delta = 0,095/0,05$ pour PAM 4/8. 104
- 5.19 Diagramme de l'œil de PAM 8. Les lignes représentent la puissance reçue : -15 dBm ((a), (b) et (c)) et -10 dBm ((d), (e) et (f)). (a) et (d) montrent signal électrique PAM à l'entrée de l'EAM ayant passé par la méthode de compensation de la distorsion. (b) et (e) montrent signal optique PAM à la sortie de l'EAM, en ayant utilisé (a) et (d) en entrée. (c) et (f) montrent signal électrique PAM à la sortie du récepteur SOA-PIN. Paramètres utilisés : $Bias = -0,3$ V, $Drive = -1,95$ V et $\Delta = 0,05$, pour les deux puissances reçues. . . . 106
- 5.20 Diagramme de l'œil de PAM 8. Les lignes représentent la puissance reçue : -15 dBm ((a), (b) et (c)) et -10 dBm ((d), (e) et (f)). (a) et (d) montrent le signal électrique PAM à l'entrée du modulateur ayant passé par la nouvelle méthode de compensation de la distorsion introduite par l'ensemble EAM + SOA-PIN. (b) et (e) montrent le signal optique PAM à la sortie de l'EAM, en ayant utilisé (a) et (d) en entrée. (c) et (f) montrent le signal électrique PAM à la sortie du récepteur SOA-PIN. Paramètres utilisés : $Bias = -0,7/-0,81$ V, $Drive = -2,1/-1,95$ V et $\Delta = 0,02/0,02$, pour des puissances reçues de -15/-10 dBm. 107
- 5.21 SER en fonction de la puissance optique reçue, pour PAM 4 (colonne de gauche) et PAM 8 (colonne de droite), pour différents débits : 14 ((a) et (b)), 28 ((c) et (d)), 56 ((f) et (g)) et 112 Gbit/s ((h) et (i)). 109
- 5.22 SER en fonction de la puissance optique reçue pour les trois meilleurs cas : PAM 4 avec un récepteur APD; PAM 4 avec un récepteur SOA-PIN; et PAM 8 avec un récepteur SOA-PIN, pour le réseau POPI+ (un EDFA) à 112 Gbit/s. 110
- 5.23 SER en fonction de la puissance optique reçue pour les trois meilleurs cas : PAM 4 - APD, PAM 4 - SOA-PIN et PAM 8 - SOA-PIN à 112 Gbit/s, pour les connexions optiques intra et inter-racks du réseau POPI+ (un EDFA). La légende « Typique » représente les résultats sans aucune variation des paramètres de $Bias$, $Drive$, Δ , G_{SOA} , NF_{SOA} , G_{EDFA} ou NF . Les légendes « $\pm 5\%$ » et « $\pm 10\%$ » indiquent que les paramètres de $Bias$, $Drive$ et Δ ont été modifiés de ± 5 ou $\pm 10\%$, respectivement. Dans tous les cas (± 5 et $\pm 10\%$) le gain et le facteur de bruit des deux amplificateurs ont été diminué et augmenté de 5 ou 10%. 111

5.24	SER en fonction de la puissance optique reçue pour les deux meilleurs cas : PAM 4 - APD et PAM 4 - SOA-PIN à 112 Gbit/s, pour les connexions du réseau POPI+ (deux EDFA). La légende « Typique » représente les résultats sans aucune variation des paramètres. Les légendes « -5% » et « -10% » indiquent des variations de -5 et -10% sur les paramètres de <i>Bias</i> , <i>Drive</i> et Δ , respectivement. Dans tous les cas le gain et le facteur de bruit des deux amplificateurs ont été diminué et augmenté de 5 ou 10%.	113
5.25	Caractérisation des filtres trapézoïdaux. (a) Allure du filtre. A_t , B_{min} , B_{max} et P correspondent aux paramètres suivants : atténuation, largeurs de bande minimale et maximale et pente, respectivement. (b) Spectre d'un signal PAM 8 superposé avec un filtre trapézoïdal.	115
5.26	Dimensionnement des réseaux (a) POPI et (b) E-POPI.	119
5.27	Dimensionnement du réseau POPI+ avec -5 et -10% de dispersion sur les paramètres de simulation.	121

LISTE DES TABLEAUX

2.1	Principales caractéristiques des sources lasers FP, VCSEL et DFB [9].	6
2.2	Principales caractéristiques des modulateurs MZM et EAM [9].	10
2.3	Principales caractéristiques des amplificateurs optiques EDFA, Raman et SOA [9], [42], [44].	14
2.4	Tendances de la performance des processeurs, le débit total du DC et la consommation d'énergie minimale des DC à partir de l'année 2012.	27
2.5	Expressions de consommation d'énergie pour les différentes architectures d'interconnexion. N , n , k et c , étant le nombre de racks, de serveurs par rack, de commutateurs d'agrégation et cœur. La puissance consommée par les différents composants présents dans chacune de ces architectures est représenté par : $P_{TRx-élec}$ (transpondeurs électriques), P_{EPS} (EPS, consommation d'énergie par port), $P_{TRx-opt}$ (transpondeurs optiques à 10 Gbit/s), P_{OCS} (OCS, consommation d'énergie de la matrice de MEMS par port), P_{WSS} (WSS), P_{WC} (convertisseurs de longueur d'onde reconfigurables), P_{O-E}/P_{E-O} (conversions O-E/E-O), P_{SDRAM} (mémoires tampon), $P_{Tx-accordable}$ (émetteurs rapidement accordables en longueur d'onde), P_{SOA} (SOA), P_{EDFA} (EDFA), P_{BPF} (filtres passe-bande optiques), P_{QM} (files d'attente), P_{BVT} (transpondeurs de largeur de bande variable), P_{VWSS} (WSS de largeur de bande variable), P_{IAS} (structure d'ordonnement intra-rack), N_p (nombre de ports de la matrice à base de MEMS utilisé par EODCN), P_{TCAM} (TCAM), P_{IES} (structure d'ordonnement inter-racks), P_{SC} (structure de contrôle), P_{ctl-ch} (canal de contrôle), $P_{ctl-plane}$ (plan de contrôle). a correspond à la probabilité de contention pour DOS (0,2).	43
2.6	Valeurs de consommation d'énergie des composants utilisés dans chaque architecture d'interconnexion.	44
3.1	Pertes des composants à considérer pour le bilan de liaison des architectures POPI et POITR options (a) et (b).	51
3.2	Comparaison des différentes options d'E-POPI.	64
4.1	Comparaison des options POPI+, T-POPI et variante de T-POPI.	80
4.2	Valeurs de sensibilité des récepteurs PIN et APD et pertes à considérer dans le bilan de liaison de l'architecture POPI+.	81

5.1	Signaux PAM 4 et 8 en entrée du générateur de signaux arbitraires (générés sous VPItrans- missionMaker TM 9.7).	87
5.2	Puissance optique nécessaire pour un SER simulé de 10^{-3} , EVM expérimentale et par simulation à cette puissance, pour des débits symbole de 2, 3, 4, 6 GBaud.	94
5.3	Paramètres considérés pour les récepteurs PIN, APD et SOA-PIN.	101
5.4	Niveaux équidistants et prédistorus des signaux PAM 4 et 8.	103
5.5	Niveaux pré-distordus des signaux PAM 8, pour compenser la distorsion introduite par l'ensemble EAM + SOA-PIN.	107
5.6	Puissance optique nécessaire pour un SER de 10^{-3} (dBm), pour les formats de modulation PAM 4 et 8, à des débits de 14, 28, 56 et 112 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN.	109
5.7	Puissance optique nécessaire pour un SER de 10^{-3} pour les connexions intra et inter-racks (un EDFA) du réseau POPI+ et pénalités de puissance introduites par l'EDFA.	110
5.8	Puissance optique nécessaire pour un SER de 10^{-3} avec une dispersion de ± 5 et $\pm 10\%$, pour les connexions intra et inter-racks (un EDFA) du réseau POPI+.	112
5.9	Puissance optique nécessaire pour un SER de 10^{-3} (dBm) avec une dispersion de -5 et -10% , pour le réseau POPI+, chaîne complète.	113
5.10	Nombre maximal de serveurs connectés par les réseaux POPI, E-POPI et POPI+, selon la taille de grille choisie. Seulement les contraintes de l'émetteur accordable en longueur d'onde et l'espacement entre canaux ont été pris en compte.	114
5.11	Pente du filtre passe-bande optique trapézoïdal (dB/nm) requise pour PAM 4 et 8 à des débits de 14, 28, 56 et 112 Gbit/s.	116
5.12	Espacement entre canaux requis (GHz) pour les meilleurs cas étudiés.	118
5.13	Nombre maximal de serveurs par rack et de racks connectés aux étages intra-rack et inter-racks, respectivement, du réseau POPI+, avec -5 et -10% de dispersion sur les paramètres de simulation.	120
5.14	Nombre maximal de serveurs par rack et de racks connectés aux étages intra-rack et inter-racks, respectivement, du réseau POPI+.	121
5.15	Nombre maximal de serveurs par rack et de racks connectés aux étages intra-rack et inter-racks, respectivement, du réseau POPI+ (le bruit d'émission spontanée amplifiée des deux EDFA est considéré).	122
B.1	Nombre maximal de serveurs connectés au réseau POPI pour les formats de modulation PAM 4 et 8 à 14 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $\text{BER} = 10^{-3}$.	146

B.2	Nombre maximal de serveurs connectés au réseau POPI pour les formats de modulation PAM 4 et 8 à 28 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $\text{BER} = 10^{-3}$.	147
B.3	Nombre maximal de serveurs connectés au réseau POPI pour les formats de modulation PAM 4 et 8 à 56 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $\text{BER} = 10^{-3}$.	148
B.4	Nombre maximal de serveurs connectés au réseau POPI pour les formats de modulation PAM 4 et 8 à 112 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $\text{BER} = 10^{-3}$.	149
B.5	Nombre maximal de serveurs connectés au réseau E-POPI pour les formats de modulation PAM 4 et 8 à 14 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $\text{BER} = 10^{-3}$.	150
B.6	Nombre maximal de serveurs connectés au réseau E-POPI pour les formats de modulation PAM 4 et 8 à 28 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $\text{BER} = 10^{-3}$.	151
B.7	Nombre maximal de serveurs connectés au réseau E-POPI pour les formats de modulation PAM 4 et 8 à 56 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $\text{BER} = 10^{-3}$.	152
B.8	Nombre maximal de serveurs connectés au réseau E-POPI pour les formats de modulation PAM 4 et 8 à 112 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $\text{BER} = 10^{-3}$.	153

INTRODUCTION GÉNÉRALE

Un centre de données (DC) est un site physique où se retrouvent plusieurs équipements connectés permettant le traitement et le stockage d'informations et l'accès à ces informations (par exemple, les bases de données dans des entreprises publiques ou privées, comme la sécurité sociale, les mutuelles, les banques, etc.).

Les DC jouent un rôle essentiel dans les télécommunications d'aujourd'hui dû à l'explosion d'applications numériques de ces dernières années, telles que les moteurs de recherche (comme *Google*, *Yahoo*, *Bing*, etc.), les réseaux sociaux (*Facebook*, *Instagram*, *Linkdin*, etc.) ou même l'accès à des fichiers stockés sur des serveurs distants accessibles via internet (*cloud computing*). Ces applications requièrent des infrastructures leur permettant d'accéder à une quantité sans cesse croissante d'information de façon rapide et sécurisée. Ces données seront traitées et transmises aussi bien à l'intérieur qu'à l'extérieur du DC, selon besoin. Une fois à l'extérieur du DC les réseaux de télécommunications prendront le relais pour que l'information sollicitée puisse arriver jusqu'au destinataire.

Les DC sont normalement organisés en serveurs informatiques (de quelques dizaines à plusieurs milliers, selon la taille du DC), qui gèrent des services de traitement de données, ayant accès à une certaine quantité d'information via des disques durs internes ou des réseaux de stockage. Les serveurs sont regroupés dans des racks ou baies. Généralement, les serveurs appartenant à une même baie partagent un système d'alimentation et de refroidissement en commun. Des connexions à l'intérieur de la baie, entre des serveurs appartenant à des baies différentes et vers l'extérieur du DC doivent être assurées par une infrastructure réseau.

La consommation d'énergie des serveurs représente la part la plus importante de celle des DC, soit 40% ; le stockage et l'infrastructure réseau consomment 37% et 23%, respectivement [1]. La totalité des centres de données dans le monde ont consommé 1,5% de l'énergie électrique globale en 2010 et une croissance exponentielle de cette consommation d'énergie est prévue avec les technologies actuelles [2]. Dès à présent, les équipements réseau pour les centres de données représentent une part identique, voire supérieure, à celle des équipements pour les réseaux d'opérateurs.

Les services du réseau sont gérés grâce à SDN (*Software-Defined Networking*), technologie qui considère une gestion globale du réseau par l'abstraction des fonctionnalités des équipements. La commande de ces équipements est séparée du traitement des données et

déportée vers un contrôleur centralisé qui dirige l'ensemble de façon cohérente.

Cependant, les applications émergentes de nos jours requièrent des DC plus puissants [3], capables de fournir des débits plus élevés avec des consommations d'énergie réduites. L'introduction des communications par fibre optique semble être la solution, car elle permet d'atteindre des débits très élevés (jusqu'à 100 voire 400 Gbit/s par canal ont été démontrés [4]-[6]) et des consommations d'énergie plutôt basses concernant les équipements réseaux par rapport aux solutions actuelles. Les communications internes au centre de données ont des caractéristiques particulières par rapport aux communications longue distance (en matière de portée, de coût et de consommation d'énergie), qui conduisent à considérer des formats de modulation spécifiques.

Les connexions optiques pour les centres de données sont récemment devenues un sujet de recherche très populaire dans la communauté scientifique des réseaux optiques et plusieurs solutions ont été proposées ces dernières années.

Notre étude s'est effectuée dans le cadre d'un projet du Labex CominLabs appelé EPOC (*Energy-Proportional and Opportunistic Computing*), en collaboration avec l'Ecole de Mines de Nantes, l'Université de Nantes, Télécom Bretagne, l'ENIB, le CNRS et l'INSA de Rennes. EPOC consiste en la gestion respectueuse de l'environnement de tâches d'un centre de données (de la couche physique jusqu'à la couche applicative). EPOC s'intéresse à des centres de données de petite et moyenne taille où la gestion des tâches va tenir compte de la demande des services et de la disponibilité des énergies renouvelables.

Le Chapitre 2 de ce manuscrit présente l'état de l'art des technologies optiques et des architectures d'interconnexion optiques pour DC, ainsi qu'une description du projet EPOC. Ce chapitre non seulement montre les solutions proposées dans la littérature, mais il présente des comparaisons entre les différentes technologies qui justifient le choix d'une architecture spécifique (dénommée POPI) basée sur des émetteurs rapidement accordables en longueur d'onde pour le réseau du projet EPOC.

Le Chapitre 3 correspond à une étude plus approfondie de l'architecture POPI, où nous proposons des améliorations de ses pertes d'insertion. Nous présentons ensuite des variantes de POPI, sous le nom de E-POPI (*Extended POPI*), utilisant plusieurs bandes de longueurs d'onde, qui permettent au moins de doubler le nombre d'entités connectées au réseau.

Le Chapitre 4 est aussi dédié à reformuler des améliorations de POPI et d'E-POPI en termes du nombre d'entités connectées. Ce chapitre propose une nouvelle famille d'architectures d'interconnexion appelée POPI+, avec deux étages hiérarchiques, permettant une augmentation quadratique du nombre d'entités connectées.

Le Chapitre 5 présente une étude approfondie sur l'emploi du format de modulation d'amplitude à plusieurs niveaux de type PAM dans les réseaux proposés dans notre travail. Nous présentons les résultats expérimentaux jusqu'à 18 Gbit/s dans le but de valider notre outil de simulation sous VPItransmissionMakerTM9.7 en termes d'amplitude du vecteur d'erreurs.

Ensuite, nous étudions par simulation des débits binaires plus élevés selon plusieurs configurations de connexions intra-DC en utilisant différents types de récepteurs. Nous appliquons la pré-compensation de non-linéarités en tenant compte de la chaîne d'émission et de réception. Nous montrons que la pré-amplification optique permet d'améliorer considérablement la photo-détection du signal. Nos résultats sont présentés en termes de taux d'erreurs de symboles. La dernière partie de ce chapitre est dédiée au dimensionnement des réseaux POPI, E-POPI et POPI+ étudiés dans les chapitres précédents à l'aide des résultats obtenus par simulation.

Enfin, le Chapitre 6 donne les conclusions et les perspectives des travaux réalisés.

ETAT DE L'ART

2.1 Introduction

Ce chapitre présente l'état de l'art des transmissions optiques et des architectures d'interconnexion optiques pour DC, ainsi que la description du projet qui motive notre recherche. Il présente les technologies disponibles au départ de cette thèse et une comparaison entre elles, en termes de performance et de consommation d'énergie.

2.2 Technologies optiques

2.2.1 Transmission sur fibre optique

Les communications sur fibres optiques sont devenues possibles principalement grâce à l'invention du laser en 1960, du laser à semi-conducteur en 1962 et à la fabrication des fibres optiques en silice avec des faibles pertes d'atténuation en 1971.

Les transmissions par fibre optique présentent une série d'avantages très importants dans le domaine des réseaux. Parmi ces avantages :

- Très faible atténuation sur une grande plage de fréquences (par exemple de 0,2 dB/km à 1,5 μm)
- Faible coût de fabrication
- Faible encombrement
- Augmentation de la distance entre répéteurs
- Insensibilité aux perturbations électromagnétiques
- Bande passante très large (≈ 12 THz) qui offre potentiellement des capacités extrêmement élevées.

Au fil des années les émetteurs, les fibres optiques, les récepteurs et d'autres composants optiques ont été améliorés, ceci a donné place à des nouvelles technologies qui ont permis d'augmenter la distance parcourue ainsi que le débit binaire. Nous décrivons certains de ces composants par la suite et nous introduisons les principales limites physiques des systèmes optiques.

2.2.1.1 Emission optique

2.2.1.1.1 Sources lasers

Les sources lasers les plus utilisées dans les systèmes de transmission optiques à haut débit sont : les lasers à cavité Fabry-Perot (FP), à cavité verticale émettant par la surface (VCSEL) et à rétroaction répartie (DFB). Ils correspondent à des diodes laser à semi-conducteurs. Parmi leurs avantages nous avons : une petite taille, une bonne qualité spectrale, une faible consommation d'énergie, une longueur d'onde ajustable dans les fenêtres de transmission optique autour de 850, 1310 et 1550 nm [7]-[8]. Dans le cadre de nos travaux nous sommes intéressés aux sources rapidement accordables en longueur d'onde, un bref état de l'art de celles-ci sera présenté dans la section suivante.

Le laser génère une onde porteuse optique qui peut être modulée directement (en modulant le courant injecté dans le laser) ou à l'aide d'un modulateur externe.

Les lasers à semi-conducteurs rajoutent du bruit d'intensité, de fréquence ou de phase. Le bruit d'intensité relatif (RIN) nous permet de mesurer le bruit d'intensité en fonction de la fréquence, il correspond au rapport entre la moyenne de la densité spectrale des fluctuations de l'intensité optique et le carré de l'intensité optique instantanée. Le bruit de phase contribue à l'élargissement du spectre émis par le laser. Par ailleurs, en cas de modulation directe, le spectre des impulsions peut être élargi par le phénomène de chirp en fréquence (variation de la fréquence instantanée par rapport à la fréquence centrale dû à un couplage entre modulation de fréquence et d'amplitude).

Dans le Tableau 2.1 nous comparons les principales caractéristiques de ces trois sources lasers [9].

Paramètre	FP	VCSEL	DFB
Rendement	0,15 W/A	0,2 W/A	0,28 W/A
RIN	-125 dB/Hz	-140 dB/Hz	-150 dB/Hz
Puissance de sortie maximale	5 mW	1,5 mW	10 mW
Courant de seuil	≈10 mA	≈6 mA	>15 mA

Tableau 2.1. Principales caractéristiques des sources lasers FP, VCSEL et DFB [9].

2.2.1.1.2 Sources lasers accordables en longueur d'onde

Les premières sources accordables en longueur d'onde de type DBR (*Distributed Bragg Reflector*) ont été développées entre 1987 et 1995 [10]-[12]. Ce type de sources est basé sur une structure ayant trois régions différentes (Cf. Figure 2.1) : une région d'amplification (zone active), une région de contrôle de phase (PC) et une région à base de réseaux périodiques gravés dans la structure semi-conductrice constituant un réflecteur de Bragg (DBR).

Le principe d'accordabilité des sources DBR est de modifier l'indice de réfraction des régions PC et DBR en injectant des courants électriques. Pour pouvoir accorder des longueurs

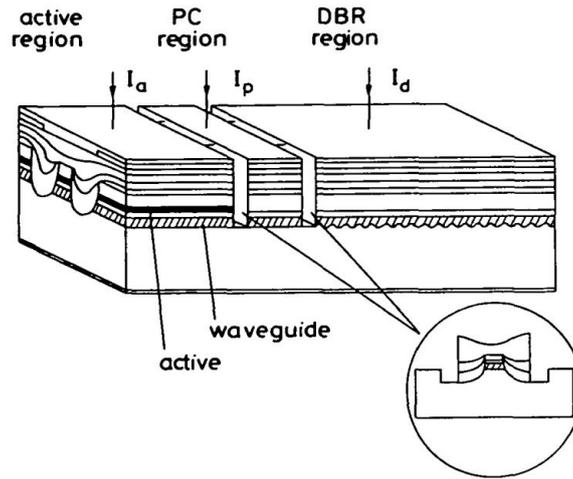


Figure 2.1. Schéma de principe des sources lasers accordables DBR [10].

d'onde en continu, ces deux courants doivent être contrôlés en même temps. La plage d'accord de ces sources est au plus de 5 à 6 nm, avec un temps d'accord minimal de 500 ps pour des longueurs d'onde consécutives [11].

Les structures DBR simples ont été améliorées à l'aide des super structures de réflexion (*SSG : Super-Structure Grating*) [13]-[18]. SSG consiste en la génération de plusieurs fréquences de réflexion utilisant un motif constitué de plusieurs réseaux de Bragg à pas variable (Cf. Figure 2.2), ce qui augmente la plage d'accord.

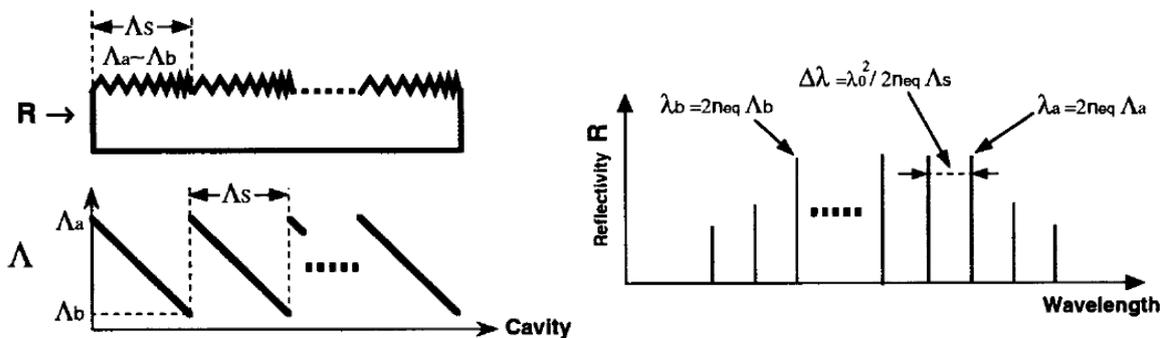


Figure 2.2. (a) Structure SSG-DBR ; (b) profil de réflectivité [13].

Les premières sources SSG-DBR ont permis d'obtenir une plage d'accord jusqu'à 83 nm, cependant, elles ne permettaient pas d'accorder la longueur d'onde en continu [13]-[14]. Des sources SSG-DBR en quasi-continu et continu ont été développées plus tard, avec des plages d'accord de 34 et 40 nm respectivement [15]-[16]. Des sources SSG-DBR plus récentes offrent des plages d'accord jusqu'à 143 nm, avec un temps d'accord de 500 ns/entre canaux voisins [17]-[18].

Il existe d'autres types de DBR avec une structure à grille échantillonnée (*SG : Sampled Grating*) [19]-[28]. La Figure 2.3 présente la structure et le profil de réflectivité des sources

SG-DBR. Ce type de source présente des plages d'accord d'environ 40 nm avec 100 canaux. La réalisation de sources SG-DBR intégrées avec un modulateur à électro-absorption avec ou sans amplificateur à semi-conducteurs a été démontrée dans [22],[23] et [24]. De plus, un système de contrôle de courant à été étudié et implémenté dans [25], qui offre plus de stabilité. Dans [28] les auteurs présentent une source SG-DBR avec un temps d'accord inférieur à 5 ns.

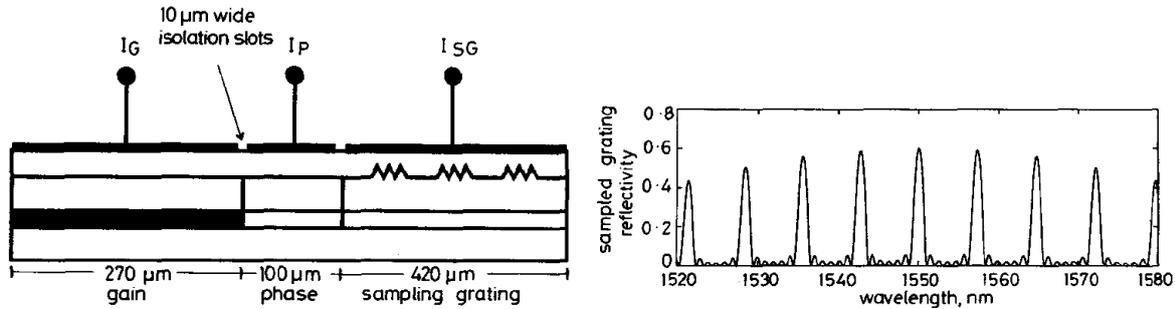


Figure 2.3. (a) Structure SG-DBR ; (b) profil de réflectivité [19].

Finalement, des sources mixtes SSG et SSG-DBR peuvent atteindre des fortes puissances (jusqu'à 45 mW) [29].

2.2.1.1.3 Modulateurs optiques

La modulation externe est utilisée par les systèmes optiques à haut débit et ou sur des longues distances afin d'éviter l'effet de chirp de fréquence ou pour atteindre des vitesses de modulation que ne permet pas la bande passante électrique du laser. Les modulateurs externes les plus utilisés dans ces systèmes sont les modulateurs Mach Zehnder (MZM) [30] et les modulateurs à électro-absorption (EAM) [31]-[32]. Nous les décrivons ci-dessous.

MZM

Le modulateur Mach Zehnder sépare la puissance du signal optique en entrée également entre deux branches. Ces deux signaux sont recombinaés après leur parcours dans ces branches. Le signal optique résultant se retrouve dans un état d'interférence constructive ou destructive partielle pour un changement de phase relatif de π entre les deux branches (voir Figure 2.4). En effet, l'indice de réfraction n de ces branches (LiNbO_3) dépend de la tension V appliquée sur ses électrodes selon :

$$\Delta_{\varphi} = \frac{2\pi}{\lambda} \Delta_n L = \pi \frac{V}{V_{\pi}} \quad (2.2.1)$$

Où Δ_{φ} et Δ_n représentent des changements de phase et d'indice de réfraction, respectivement. λ , L et V_{π} correspondent à la longueur d'onde du signal en entrée du modulateur, à la longueur des branches et à la tension de polarisation pour avoir un déphasage de π . Avec des tensions de polarisation $V_1 = -V_2 = V$, appliquées à chaque électrode, le champ électrique en sortie $E_s(t)$ est décrit en fonction du champ électrique en entrée $E_e(t)$ par :

$$E_s(t) = \cos\left(\frac{\pi V + V_{DC}}{2} \frac{V}{V_\pi}\right) E_e(t) \quad (2.2.2)$$

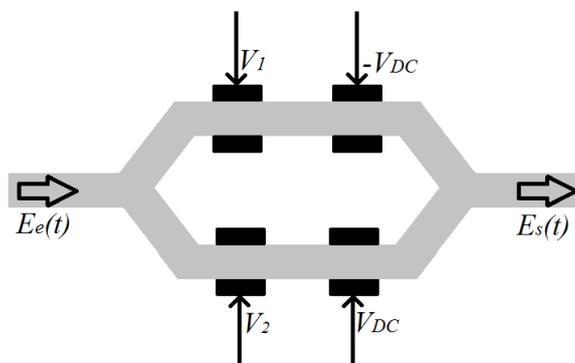


Figure 2.4. Schéma de principe du modulateur Mach Zehnder.

Ce modulateur peut être utilisé pour générer des formats de modulation en amplitude et en phase.

EAM

Le modulateur à électro-absorption utilise généralement l'effet « Franz-Keldysh » dans des matériaux semi-conducteurs. La jonction p-n est transparente lorsque la tension de polarisation est nulle et elle devient opaque pour des tensions élevées (Cf. Figure 2.5). Ce dispositif peut être facilement intégré à une source laser.

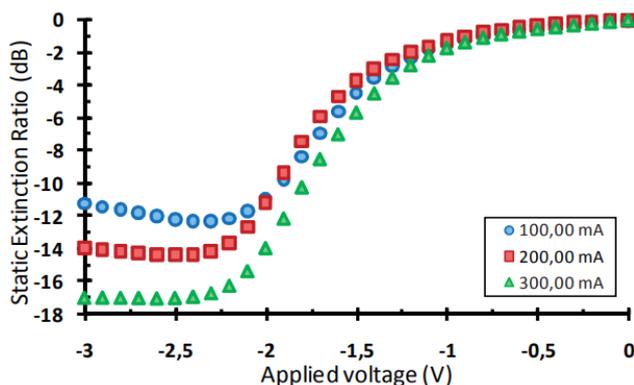


Figure 2.5. Taux d'extinction de l'EAM [33].

Ce modulateur peut être utilisé pour générer seulement des formats de modulation en amplitude.

Nous comparons à l'aide du Tableau 2.2 les principales caractéristiques de ces deux modulateurs optiques [9].

Paramètre	MZM	EAM
Degré d'intégration	LiNbO ₃ faible	élevé [33]-[34]
Taux extinction	jusqu'à 25 dB	≈10 dB
Tension	4-5 V	1-3 V
Bande passante électrique	≈40 GHz	>60 GHz

Tableau 2.2. Principales caractéristiques des modulateurs MZM et EAM [9].

2.2.1.2 Fibre optique

La fibre optique est un guide d'onde qui exploite les propriétés de réfraction de la lumière. Elle est constituée d'un cœur entouré d'une gaine. Le cœur de la fibre a un indice de réfraction légèrement plus élevé que celui de la gaine, ceci permet de confiner la lumière qui se trouve entièrement réfléchi de multiples fois à l'interface entre les deux matériaux. L'ensemble est recouvert par une gaine plastique de protection. Il existe deux types de fibres optiques : la fibre multimode et la fibre monomode, nous les décrivons ci-dessous.

2.2.1.2.1 Fibres optiques multimode et monomode

Les fibres multimodes (MMF), transportent plusieurs trajets lumineux appelés modes. Le fait de transmettre plusieurs modes provoque un étalement temporel du signal proportionnel à la longueur de la fibre (dispersion inter-modale). Ce qui restreint l'utilisation de cette fibre à des connexions à bas débits et à des faibles distances. Toutefois, la dispersion modale peut être diminuée en raccourcissant la distance ou en réalisant un gradient d'indice dans le cœur de la fibre. Généralement les fibres multimodes utilisent des diamètres de cœur de 50 ou 62,5 μm . Quelques fibres multimodes plus récentes, de type OM4, permettent d'atteindre 1 Gbit/s sur des distances de l'ordre du km.

Les fibres monomodes (SMF) ont été conçues avec des diamètres de cœur plus petits (environ 10 μm) afin de ne pouvoir transporter qu'un seul mode (ou plus précisément deux modes de polarisation orthogonaux). Les fibres monomodes permettent la transmission optique à très haut débits et sur de très longues distances (par exemple, pour les câbles sous-marins des lignes intercontinentales).

2.2.1.2.2 Dégradations résultants de la propagation optique

Nous pouvons classer les effets contraignants à la propagation du signal optique dans la fibre comme linéaires (atténuation, dispersions chromatique et par mode de polarisation), non linéaires dus à l'effet Kerr (SPM, XPM et FWM) et dus à l'effet de diffusion stimulée (Brillouin et Raman). Nous les décrivons par la suite.

Atténuation

L'atténuation α (dB/km) correspond aux pertes de puissance du signal optique au cours de sa propagation dans la fibre optique. Elle est donnée par :

$$\alpha = \frac{10}{L} \log \left(\frac{P_{Rx}}{P_{Tx}} \right) \quad (2.2.3)$$

Où L est la longueur de la fibre optique. P_{Rx} et P_{Tx} représentent la puissance optique du signal à la sortie et à l'entrée de la fibre, respectivement.

Les fibres monomodes fabriquées en silice sont les plus utilisées pour les transmissions à haut débit. Elles présentent des valeurs d'atténuation qui varient en fonction de la longueur d'onde selon la Figure 2.6 [35].

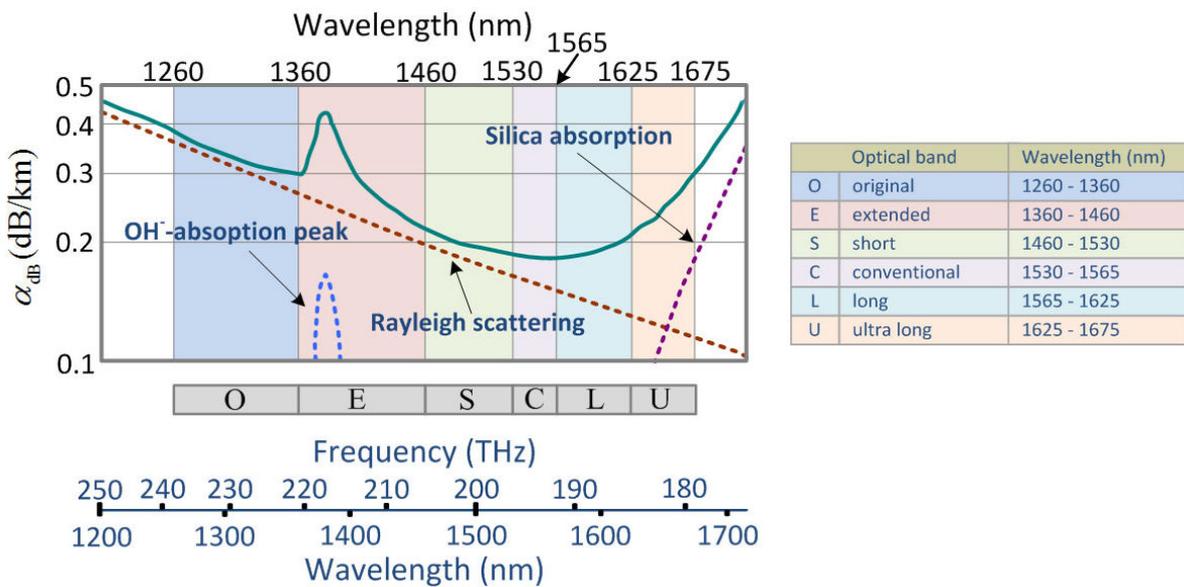


Figure 2.6. Atténuation en fonction de la longueur d'onde [35].

Plusieurs bandes de transmission optiques ont été définies par l'ITU : O, E, S, C, L et U, à partir de ce spectre de pertes. La bande C présente les pertes les moins élevées (0,2 dB/km) et elle est aujourd'hui la plus utilisée pour les transmissions optiques à haut débit sur des longues distances.

Dispersion chromatique

La dispersion chromatique est caractérisée par l'étalement temporel du signal optique dû au fait que l'impulsion optique est composée par plusieurs longueurs d'onde qui se propagent à des vitesses différentes. La dispersion chromatique dépend des caractéristiques physiques propres au matériau et au profil d'indice de la fibre. Par exemple, pour une fibre de type G652, la dispersion chromatique est minimale vers 1300-1310 nm et de l'ordre de 17 ps/(nm km) autour de 1550 nm. Lorsque la dispersion chromatique (retard τ_g) accumulée (dans une longueur L de fibre) provoque des impulsions plus larges que le temps de symbole (défini par le débit binaire), elle induit des erreurs binaires à la détection. Ce retard est donné par :

$$\tau_g = -L \frac{\lambda^2}{2\pi c} \frac{d\beta}{d\lambda} \quad (2.2.4)$$

Où β est la constante de propagation : $\beta = n_{eff}(\omega/c)$. Avec n_{eff} , l'indice de réfraction effectif et ω , la pulsation du signal [36]. L'élargissement de l'impulsion $\Delta\tau_g$ est donné par :

$$\Delta\tau_g = LD\Delta\lambda \quad (2.2.5)$$

Avec $\Delta\lambda$, la largeur spectrale du signal et D paramètre de dispersion en ps/(nm km) :

$$D = -\frac{2\pi c}{\lambda^2} \frac{d^2\beta}{d\omega^2} \quad (2.2.6)$$

Dispersion modale de polarisation (PMD)

La dispersion modale de polarisation est un élargissement temporel de l'impulsion optique, elle peut être observée entre deux signaux à la même longueur d'onde avec des modes de polarisation orthogonaux. Ces deux composantes vont se propager à des vitesses différentes dues à la biréfringence de la fibre optique. Le retard temporel ΔT entre ces deux signaux est décrit par :

$$\Delta T = L|\beta_x - \beta_y| = D_p\sqrt{L} \quad (2.2.7)$$

Où D_p correspond au paramètre de dispersion modale de polarisation en ps/km^{-1/2}.

L'effet Kerr correspond à une dépendance de l'indice de réfraction en fonction de l'amplitude du signal optique. Cet effet peut être observé à des puissance optiques élevées.

Auto-modulation de phase (SPM) et modulation de phase croisée (XPM)

Pour les transmissions mono-canal à forte puissance d'émission, l'effet Kerr induit un déphasage de l'impulsion optique, appelé auto-modulation de phase, qui modifie son spectre et génère du chirp temporel. Pour les transmissions à plusieurs canaux, le déphasage de l'impulsion va non seulement dépendre de la puissance optique du canal, mais aussi de celle des autres canaux.

Mélange à quatre ondes (FWM)

Le mélange à quatre ondes consiste en la génération d'un signal optique à partir de trois autres signaux optiques à fréquences différentes. Le signal généré est cause de gêne majeure par diaphonie pour les transmissions optiques par multiplexage en longueur d'onde.

Effet de diffusion stimulée Brillouin et Raman

Il s'agit d'un transfert d'énergie d'un photon vers d'autres photons d'énergie moins élevée, la différence d'énergie est diffusée sous forme de phonon. Le phonon diffusé par l'effet Brillouin est de nature acoustique et celui diffusé par l'effet Raman, de nature optique. Dans les deux cas, la puissance optique du photon incident aura diminuée.

La diffusion stimulée est une cause de gêne pour les transmissions optiques à plusieurs canaux, car les basses fréquences vont être amplifiées par les hautes fréquences quand la puissance d'émission totale dépasse le seuil qui déclenche cet effet. Cependant, la diffusion stimulée peut être exploitée pour faire de l'amplification optique de type Brillouin, Raman distribué et Raman localisée.

2.2.1.3 Amplificateurs optiques

Le rôle principal des amplificateurs optiques est de compenser l'atténuation de la fibre pour ainsi augmenter la longueur des liens optiques. Avant l'apparition de l'amplification optique des conversions intermédiaires entre les domaines optique, électrique et optique (O-E-O) étaient nécessaires pour atteindre des longues distances. Ce genre de conversion O-E-O, présente des coût et consommation d'énergie élevés. Nous décrivons par la suite les amplificateurs les plus utilisés : EDFA, Raman et SOA.

Le développement des fibres dopées à l'Erbium pendant les années 80 a motivé fortement la recherche des systèmes de transmission optiques utilisant des fibres à régions actives dans la bande C. Les premiers amplificateurs à fibre dopée à l'Erbium (EDFA) ont été présentés en 1987 [37]-[38]. Ces amplificateurs utilisent une pompe optique qui va déclencher l'émission stimulée des photons identiques au signal de transmission. Ils présentent des forts gains ($30\text{-}50\text{ dB}^1$) et des puissances de saturation élevées. Ces amplificateurs sont aujourd'hui les plus utilisés pour des transmissions longue distance (liens transocéaniques). Un des principaux inconvénients est leur bande spectrale utile de l'ordre de 30 nm généralement dans la bande C. Cependant, des EDFA en bande L ont été développés [40].

Les amplificateurs Raman distribués utilisent la fibre de transmission pour amplifier le signal optique grâce à l'effet de diffusion stimulée. Le fait d'amplifier le signal au fur et à mesure de sa propagation limite la puissance optique instantanée, et par conséquent les effets non-linéaires. La largeur de bande des amplificateur Raman et leur bande d'amplification varient selon la quantité et la position des pompes optiques utilisées. Plusieurs études ont montré qu'il est possible d'optimiser les pompes afin d'avoir un gain plat sur des bandes très larges [41]. L'inconvénient principal de l'amplification Raman est sa faible efficacité énergétique due au fait que la fibre utilisée pour l'amplification n'a pas de région active. Il n'est pas possible donc de générer des fort gains pour des faibles longueurs de fibre. Nonobstant, leur efficacité énergétique peut être améliorée en les associant avec des EDFA [42]-[43].

Les amplificateurs à semi-conducteurs (SOA) sont des dispositifs opto-électroniques. Ils fournissent un gain optique lorsque leur région active est alimentée à l'aide d'un courant électrique. Ces amplificateurs présentent des gains et des puissances de saturation inférieurs

1. Un gain de 50 dB à été démontré expérimentalement dans [39]. Des gains aussi élevés ne sont pas forcément désirables, le rapport signal à bruit peut être faible et de fortes puissances peuvent induire des effets non-linéaires.

à ceux des EDFA. Par contre, ils ont l'avantage d'offrir une bande d'amplification plus élevée que celle de l'EDFA et la possibilité d'être intégrés à d'autres composants opto-électroniques (par exemple pour augmenter la sensibilité des récepteurs). En outre, les SOA peuvent être utilisés pour la réalisation de portes logiques tout-optiques et de convertisseurs de longueur d'onde.

Ces trois types d'amplificateurs optiques rajoutent une certaine quantité de bruit au signal amplifié dû à l'émission spontanée des photons et à leur éventuelle amplification. Ce bruit, appelé bruit d'émission spontanée amplifiée (ASE), va dépendre du gain optique ainsi que de la bande passante de l'amplificateur. En pratique, le bruit ASE est représenté par un facteur de bruit (NF), qui peut être exprimé en fonction du rapport signal sur bruit optique (OSNR) avant et après chaque amplification selon :

$$OSNR_s = OSNR_e - NF, \quad NF = 10 \log \left(\frac{2\eta_{sp}(G-1)}{G} \right) \quad (2.2.8)$$

Où $OSNR_s$ et $OSNR_e$ représentent l'OSNR (en dB) en sortie et en entrée du lien contenant un amplificateur optique. G est le gain et η_{sp} correspond au facteur d'inversion de population.

Finalement, nous comparons ces trois types d'amplificateurs à l'aide du Tableau 2.3.

Paramètre	EDFA	Raman	SOA
Gain optique	30-50 dB	20-30 dB	20-35 dB
Bande optique	1530-1560 nm	variable ^a	1280-1650 nm
Bande spectrale utile	30 nm	variable ^b	30-100 nm
Puissance de saturation	10-25 dBm	variable	5-25 dBm
Facteur de bruit	3-7 dB	2-5 dB	5,5-7 dB
Temps de récupération du gain	<10 ms	2-10 ns	<1 ns
Longueur de fibre d'amplification	10-20 m	> 40 km	-
Possibilité d'intégration	non	non	oui
Pompage	optique	optique	électrique

^aselon les longueurs d'onde des pompes

^bselon le nombre de pompes

Tableau 2.3. Principales caractéristiques des amplificateurs optiques EDFA, Raman et SOA [9], [42], [44].

2.2.1.4 Réception optique

Les récepteurs optiques ont pour but de convertir la puissance optique reçue en courant électrique pour ensuite reconstituer la séquence binaire envoyée. Dans cette section nous présentons les trois types de récepteurs optiques que nous serons amenés à considérer par la suite : photodiodes PIN et APD et l'ensemble SOA + photodiode PIN.

2.2.1.4.1 Photodiode PIN

La photodiode PIN est une structure constituée d'une couche de matériau intrinsèque i entre deux couches de matériaux dopés p et n . Une polarisation inverse appliquée sur la jonction donne naissance à un champ électrique dans la zone intrinsèque. Lorsqu'un photon d'énergie suffisante est absorbé par cette zone, il génère une paire électron/trou qui sera dissociée sous les effets du champ électrique, ses photo-porteurs seront déplacés vers la région dans laquelle ce dernier est majoritaire. Le déplacement des photo-porteurs est à l'origine du photo-courant. Le photo-courant I_{PIN} étant donné par :

$$I_{PIN} = RP_{Rx} \quad R = \eta \frac{\lambda q}{hc} \quad (2.2.9)$$

Où R et P_{Rx} représentent la détectivité du récepteur (A/W) et la puissance optique reçue (W), respectivement. η , λ , q , h et c correspondent au rendement quantique de la photodiode, la longueur d'onde du signal optique, la charge ($1,6 \times 10^{-19}$ C), la constante de Planck ($6,6 \times 10^{-34}$ m²kg/s) et la célérité de la lumière (3×10^8 m/s), respectivement.

Le bruit électrique de la photodiode PIN est représenté par la somme des bruits quadratiques de grenaille σ_{gr} et thermique σ_{th} selon : $\sigma = \sqrt{\sigma_{gr}^2 + \sigma_{th}^2}$ [36]. Le premier est donné par : $\sigma_{gr}^2 = 2qRP_{Rx}B_e$. Le deuxième, par : $\sigma_{th}^2 = \frac{4kTB_e}{R_L}$, avec k constante de Boltzmann ($1,38 \times 10^{-23}$ m²kg/(s²K)) ; T , la température en K ; B_e , la bande passante du récepteur (Hz) et R_L (Ω), la résistance de charge. Le bruit thermique est prépondérant pour de faibles puissances en émission.

2.2.1.4.2 Photodiode à avalanche (APD)

La photodiode APD rajoute une couche à la structure de la photodiode PIN. Cette dernière va stimuler l'effet d'avalanche qui multiplie le nombre de paires trou-électron. Autrement dit, le photo-courant de la photodiode APD I_{APD} est multiplié d'un facteur M selon :

$$I_{APD} = MRP_{Rx} \quad (2.2.10)$$

Le facteur M dépendra des caractéristiques physiques de la couche rajoutée, en général $M = 10$ pour des photodiodes APD *InGaAs*.

Le bruit électrique de la photodiode ADP s'exprime toujours sous la forme : $\sigma = \sqrt{\sigma_{gr}^2 + \sigma_{th}^2}$, mais le bruit de grenaille σ_{gr} intègre maintenant le facteur de bruit de l'effet d'avalanche F_A qui dépend du coefficient de ionisation κ_A :

$$\sigma_{gr}^2 = 2qM^2F_ARP_{Rx}B_e, \quad F_A = \kappa_A M + (1 - \kappa_A) \left(2 - \frac{1}{M}\right) \quad (2.2.11)$$

2.2.1.4.3 Photodiode PIN + pré-amplification SOA (SOA-PIN)

La sensibilité de la photodiode PIN est améliorée grâce à la pré-amplification optique avec l'utilisation d'un amplificateur à semi-conducteur qui peut être intégré monolithiquement [45]. Ce type de récepteur peut être utilisé dans les transmissions optiques à haut débit et sur des longues et courtes distances [46]-[48]. Parmi ses avantages nous signalons des faibles pénalités par dépendance au mode de polarisation (<1 dB), un facteur de bruit inférieur à 8,5 dB et une largeur de bande électrique à -3 dB d'environ 50 GHz [46].

La détectivité du SOA-PIN $R_{SOA-PIN}$ va dépendre du gain de l'amplificateur G_{SOA} et de la détectivité de la photodiode PIN R selon :

$$R_{SOA-PIN} = G_{SOA} R P_{Rx} \quad G_{SOA} = \frac{P_{s,SOA}}{P_{e,SOA}} \quad (2.2.12)$$

Avec $P_{s,SOA}$ et $P_{e,SOA}$, puissance en sortie et entrée du SOA, respectivement.

Le bruit électrique du récepteur SOA-PIN est représenté par la somme des bruits quadratiques de grenaille σ_{gr} , thermique σ_{th} , de battement ASE-ASE $\sigma_{ASE-ASE}$ et de battement signal-ASE σ_{s-ASE} selon :

$$\sigma = \sqrt{\sigma_{gr}^2 + \sigma_{th}^2 + \sigma_{ASE-ASE}^2 + \sigma_{s-ASE}^2}$$

Ces dernières sources de bruit correspondent à la détection du bruit d'émission spontanée amplifiée du SOA par la photodiode PIN. Le bruit ASE va de même contribuer au bruit de grenaille du SOA-PIN en accord avec :

$$\sigma_{gr}^2 = 2qR(G_{SOA}P_{Rx} + P_{sp})B_e, \quad S_{sp}(\nu) = (G_{SOA} - 1)\eta_{sp}h\nu \quad (2.2.13)$$

Où S_{sp} et η_{sp} correspondent à la densité spectrale et le facteur de bruit de l'émission spontanée du SOA.

Les bruits de battement ASE-ASE et signal-ASE sont donnés par :

$$\sigma_{ASE-ASE}^2 = 4R^2 S_{sp}^2 \Delta\nu_{sp} B_e \quad (2.2.14)$$

$$\sigma_{s-ASE}^2 = 4R^2 G_{SOA} P_{Rx} S_{sp} B_e \quad (2.2.15)$$

Avec $\Delta\nu_{sp}$: largeur du spectre mesuré de l'émission spontanée amplifiée.

2.2.1.5 Systèmes de transmission WDM

Le multiplexage en longueur d'onde consiste en la transmission optique de plusieurs canaux, chacun à une longueur d'onde donnée (Figure 2.7). L'ensemble des canaux est injecté dans une fibre optique unique à l'aide d'un multiplexeur ou d'un coupleur. A la réception, les

canaux sont séparés dans des fibres optiques différentes grâce à un démultiplexeur ou un coupleur suivi de filtres optiques, dans le but d'être détectés par des récepteurs différents. Le multiplexage par longueur d'onde permet notamment d'augmenter le débit binaire total dans la fibre optique.

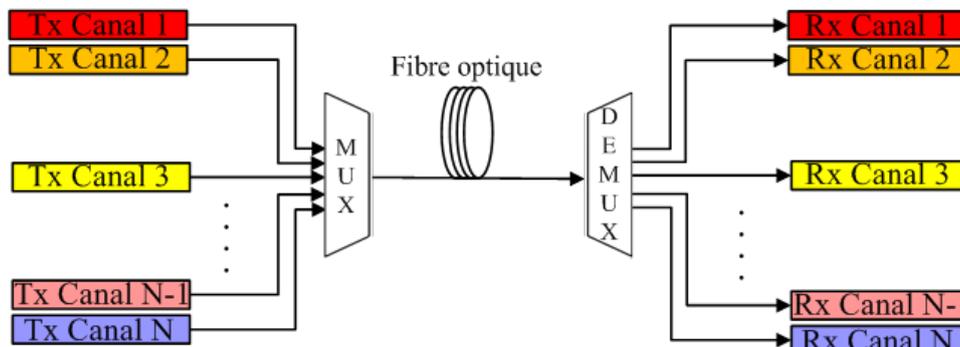


Figure 2.7. Système de transmission WDM à N canaux. MUX : multiplexeur, DEMUX : démultiplexeur.

Les systèmes de transmission WDM ont évolué en agissant sur trois aspects différents : en augmentant le débit binaire par canal (de 2,5 Gbit/s en 1995 jusqu'à 100 Gbit/s en 2015) ; en réduisant l'espacement entre canaux (de 100 à 50 ou 25 GHz) grâce l'évolution des filtres optiques ; et en élargissant la bande optique exploitée (extension de la bande de transmission C vers les bandes S et L).

2.2.1.6 Formats de modulation

Le format de modulation correspond à la façon dont l'information est codée afin de pouvoir la transmettre dans un milieu en particulier. C'est-à-dire, le signal contenant l'information à transmettre (onde modulante) est transportée par une onde électromagnétique ou mécanique (onde porteuse) par le biais d'une modification d'amplitude, de fréquence ou de phase de l'onde porteuse.

Les formats de modulation le plus simples dans les communications analogiques et numériques sont basées sur la modulation d'amplitude : AM en analogique, ASK (*Amplitude-Shift Keying*) ou OOK (*On-Off Keying*) en numérique. En optique, les formats de modulation les plus utilisés sont RZ (*Return to Zero*) et NRZ (*Non Return to Zero*).

La méthode de détection la plus utilisée actuellement pour les formats de modulation d'amplitude est la détection directe d'intensité à travers la mesure de courant électrique aux bornes d'une photodiode.

Le débit a augmenté jusqu'à 10 voire 40 Gbit/s en gardant un format de modulation à deux niveaux et avec le même espacement entre canaux WDM. Les grilles classiques et régulières au pas de 50 ou 100 GHz étaient suffisamment larges par rapport à la largeur de

bande du signal. Par contre, il n'est pas possible de faire passer un signal à 100 Gbit/s avec une modulation à deux états dans un canal de 50 GHz. Par exemple, le premier lobe d'un signal NRZ à 100 Gbit/s a une largeur de 200 GHz, mais sachant qu'il suffit de prendre en compte 75% de cette largeur, une largeur de bande en réception 150 GHz est nécessaire pour détecter ce signal. On voit bien qu'il est nécessaire d'améliorer l'efficacité spectrale afin de pouvoir augmenter le débit sans devoir augmenter l'espacement entre canaux.

Des formats de modulation d'amplitude (PAM M) [49], de phase (PSK M) [50]-[51] et d'amplitude et phase en quadrature (QAM M) [52] à M états permettent d'associer plusieurs bits par symbole, ce qui réduit la largeur de bande du signal à $1,5 \times$ débit symbole à débit constant. Cependant, le fait d'augmenter le nombre de bits par symbole va réduire l'écart entre chaque symbole et donc rendre le signal plus sensible au bruit. Une façon de faire face à ce bruit est d'augmenter la puissance optique à l'émission. Nonobstant, il faut éviter de ne pas dépasser les seuils de puissance qui vont favoriser l'apparition des effets non-linéaires.

Les formats de modulation PAM M peuvent être implémentés avec de la détection directe aussi bien qu'avec de la détection cohérente. Cette dernière permet d'avoir des sensibilités de réception améliorées, car ce type de détection retrouve l'intensité et la phase du signal optique en réception à l'aide d'un oscillateur local. Les formats PSK et QAM doivent être implémentés avec une détection cohérente.

Il existe d'autres formats de modulation comme les modulations multiporteuses. Le plus connu étant le format de modulation par multiplexage de fréquences orthogonales (OFDM). L'OFDM peut être implémenté avec de la détection directe [53]-[54], surtout pour les transmissions courte distance, ainsi qu'avec de la détection cohérente [55]-[56], pour les transmissions longue distance.

En outre, le multiplexage de deux modes de polarisation orthogonaux (PDM) permet de doubler les débits atteints par les formats de modulation décrits ci-dessus.

Les systèmes de transmission WDM à grille fixe sont en train de migrer vers des grilles flexibles qui adaptent l'espacement entre canaux en fonction de la largeur de bande du signal optique. Ceci permet d'augmenter l'efficacité spectrale de la bande utilisée. Selon les recommandations de l'ITU pour les transmissions à grille flexible, les fréquences centrales des canaux doivent être placées sur une grille au pas de 6,25 GHz et avec une largeur de canal égale à un multiple entier de 12,5 GHz.

2.2.2 Réseau et commutation optiques

2.2.2.1 Réseaux optiques

Les systèmes de transmission WDM sont utilisés dans les cœurs des réseaux d'opérateurs qui relient des grandes villes avec des systèmes de transmission longues distances. Ces réseaux doivent transporter les trafics insérés et extraits dans les nœuds situés dans les principales agglomérations.

2.2.2.1.1 Réseau opaque

Il s'agit de la transmission optique des canaux WDM point-à-point. C'est-à-dire, la transmission de données entre chaque nœud du réseau est optique, cependant au niveau du nœud tous les canaux doivent être traités dans le domaine électrique, qu'ils soient à destination ou en provenance du nœud ou simplement en transit. Cela nécessite beaucoup de conversions O-E-O, qui sont chères et dont la consommation d'énergie est élevée. L'insertion/extraction du trafic dans les nœuds électriques se fait soit sous forme de circuits multiplexés temporellement dans chaque canal optique ou, de plus en plus, sous forme de paquets.

2.2.2.1.2 Réseau transparent en mode circuit

Dans ces réseaux seuls les canaux insérés et extraits sont traités dans une matrice électrique. Les canaux en transit ne subissent pas de conversion O-E-O. Pour des raisons de flexibilité (restauration en cas de panne, évolution du trafic, etc.) l'ensemble des canaux traverse une matrice de commutation de circuits optiques (OCS). Pour cette matrice les temps de commutation de quelques millisecondes sont largement suffisants. Dans ces conditions, les canaux restent sous forme optique depuis la source jusqu'à leur destination.

Il existe des cas intermédiaires entre les réseaux opaques et transparents, dénommés réseaux translucides. Dans ce type de réseau certains canaux subissent des conversions O-E-O dans un ou plusieurs des nœuds intermédiaires.

2.2.2.1.3 Réseau transparent en mode paquet

Depuis plusieurs années plusieurs laboratoires ont proposé de faire jouer un plus grand rôle aux technologies optiques, en les utilisant pour commuter des paquets ou des groupes de paquets sous forme optique. L'idée étant de concilier les avantages en matière de consommation d'énergie de la commutation optique avec le partage efficace des ressources de transmission permis par le mode paquet. On parle ainsi de commutation de paquets optiques (OPS).

La commutation de rafales optiques [57], quant à elle, est une technique intermédiaire. L'OBS regroupe un grand nombre de paquets qui vont constituer la rafale. La taille (et donc la durée) de la rafale peut être fixe ou flexible. L'entête de la rafale, indiquant son début et sa durée, est transmise séparément dans un canal à part qui sera traité dans le domaine électrique.

La principale difficulté pour l'implémentation de ces techniques est qu'elles nécessitent des dispositifs de commutation rapide dont les temps de commutation sont de l'ordre de la dizaine de nanosecondes pour l'OPS et de quelques microsecondes pour l'OBS.

2.2.3 Autres technologies optiques

2.2.2.1 Dispositifs passifs

- **Coupleur** : dispositif qui sert à distribuer la puissance du signal optique d'une fibre optique vers deux ou plusieurs fibres optiques. Le coupleur peut également faire le processus inverse, c'est-à-dire, combiner le signal optique de deux ou plusieurs fibres optiques en une seule.
- **Circulateur** : dispositif optique à trois ports qui permet d'orienter la sortie du signal en fonction du port d'entrée et de sa direction de propagation.
- **Filtre passe-bande optique (BPF)** : dispositif optique qui permet de laisser passer une bande spectrale particulière et de bloquer les longueurs d'onde autour de cette région.
- **Réseau de Bragg (FBG)** : structure dans laquelle des couches de deux matériaux à indices de réfraction différents sont alternées périodiquement de telle sorte de provoquer la réflexion d'une longueur d'onde en particulier et de laisser passer les autres. En association avec un circulateur il peut être utilisé comme filtre passe-bande et coupe-bande en même temps.
- **(Dé)multiplexeur WDM** : le démultiplexeur est un dispositif qui permet de séparer ℓ longueurs d'onde présentes dans la fibre optique vers ℓ fibres optiques différentes. Le multiplexeur fait le processus inverse, il combine plusieurs longueurs d'onde (une longueur d'onde par fibre optique) sur une seule fibre optique.
- **Duplexeur** : dispositif qui permet de séparer deux plages de longueurs d'onde appartenant à différentes bandes de transmission optique d'une fibre optique parmi deux fibres optiques différentes. Il permet également de faire le processus inverse, il combine deux fibres (une plage de longueurs d'onde par fibre optique) en une seule fibre optique.
- **AWG (*Arrayed-Waveguide Grating*)** : dispositif qui permet l'acheminement de chaque longueur d'onde vers une sortie différente (la longueur d'onde w de l'entrée i est acheminée vers la sortie $[(i + w - 2) \bmod \ell] + 1$, ou ℓ est le nombre de ports).

2.2.3.2 Dispositifs actifs

- **Filtre passe-bande optique reconfigurable** : dispositif optique qui permet de laisser passer une bande spectrale particulière et de bloquer les longueurs d'onde autour de cette région. La bande peut être reconfigurable selon besoin.
- **Commutateur sélectif par longueur d'onde (WSS)** : dispositif $1 : \ell$ qui peut partitionner l'ensemble entrant des longueurs d'onde vers différents ports. Chaque longueur d'onde ou groupe de longueurs d'onde peut être assigné pour être ache-

miné vers un port différent. Ce dispositif est utilisé pour la commutation de circuits optiques.

- **Commutateur à base de systèmes micro-électromécaniques (MEMS)** : dispositif basé sur le contrôle physique d'un réseau de miroirs qui redirigent un faisceau laser en entrée afin d'établir une connexion avec la sortie désirée. Le temps de configuration est de l'ordre de la milliseconde. Ce dispositif est utilisé pour la commutation de circuits optiques.
- **Convertisseur de longueur d'onde optique** : dispositif qui permet de changer la longueur d'onde d'un signal optique en entrée sans affecter l'information transmise. Ce dispositif comprend un laser accordable en longueur d'onde, un SOA et un interféromètre Mach-Zehnder (MZI). Son temps de configuration est de l'ordre de la nanoseconde.

2.2.4 Formats de modulation pour DC

Récemment, plusieurs auteurs ont travaillé sur l'utilisation des formats de modulation multiniveaux pour les connexions optiques internes d'un DC [58]-[67]. Le défi est d'atteindre des débits élevés (>40 Gbit/s) à des bas prix et avec des faibles consommations d'énergie. Ceci exclut le format de modulation NRZ par sa faible efficacité spectrale, ainsi que les formats de modulation basés sur une détection cohérente. En conséquence, les formats de modulation les plus utilisés dans ce domaine sont : PAM, CAP (*Carrierless Amplitude and Phase*)² et OFDM avec détection directe. D'une part, PAM à 4 ou 8 niveaux augmente l'efficacité spectrale de NRZ, mais en moindre mesure que CAP ou OFDM. D'autre part, l'OFDM requiert du traitement de signal qui est cher et qui augmente la consommation d'énergie du DC [68].

J. L. Wei et al. ont présenté une étude approfondie des technologies nécessaires pour la mise en œuvre de ces formats [59]. Cette étude inclut des comparaisons de consommation d'énergie entre ces formats (Cf. Figure 2.8). Les formats PAM 4, CAP 16 et OFDM (en combinaison avec QAM 16), à débit constant, consomment 1,5, 2 et 4 fois plus que NRZ.

Au vu de ces résultats, le format PAM semble être le format de modulation le plus adapté pour les DC. Pour cette raison, nous l'étudions en détail dans cette section.

2.2.4.1 Format de modulation PAM

La modulation par amplitude d'impulsions PAM est une technique de modulation exclusivement par amplitude (phase et fréquence restent fixes). En fonction du nombre de symboles de modulation possibles, nous parlons de PAM M (par exemple PAM 2 (équivalent

2. Le format de modulation CAP utilise deux signaux PAM dé-corrélés de 90° grâce à des filtres adaptés. Il offre la même efficacité spectrale que le format QAM, mais avec une mise en place plus simple et qui consomme moins d'énergie [61]-[63].

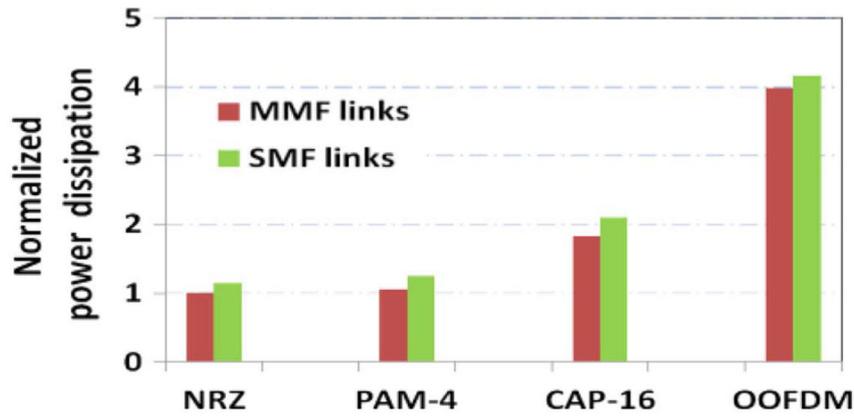


Figure 2.8. Consommation d'énergie des formats de modulation NRZ, PAM, CAP et OFDM, à un débit binaire constant de 28 Gbit/s [59].

au format NRZ), PAM 4, PAM 8, etc). Le choix du nombre de symboles de modulation possible dépend de l'immunité du signal par rapport au bruit (distance entre les points) ou de l'énergie par bit (distance à l'origine).

La génération du signal PAM consiste à passer d'un signal numérique binaire (contenant l'information à transmettre) à un signal numérique ayant M niveaux d'amplitude. Plus concrètement, nous considérons le système de transmission optique présenté sur le schéma de la Figure 2.9 ((a) émission (Tx), (b) réception (Rx)).

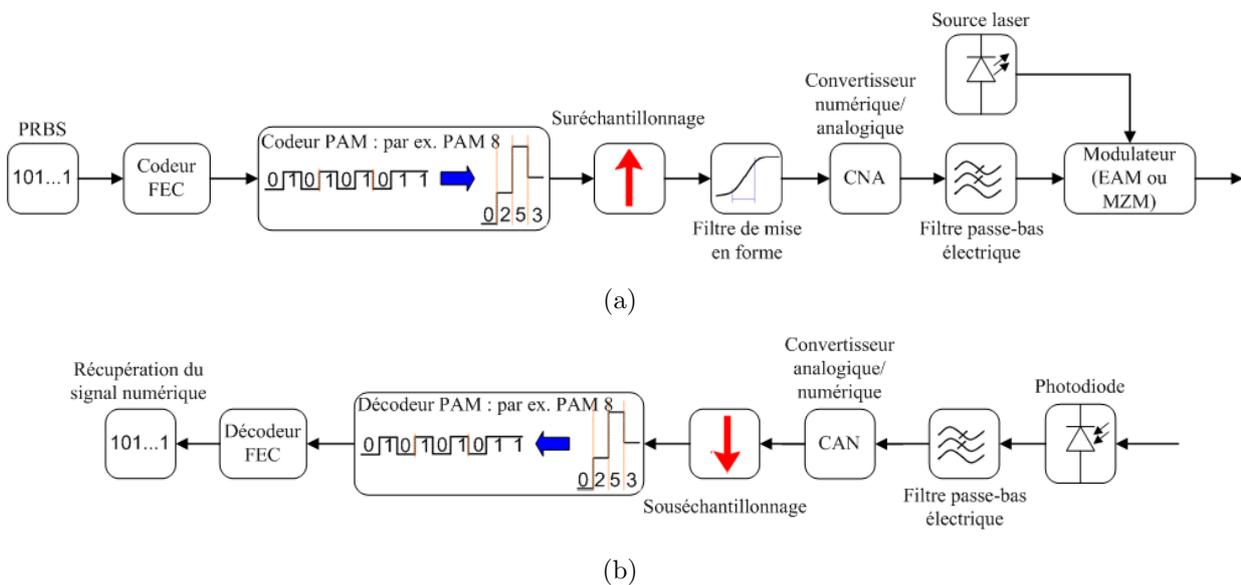


Figure 2.9. Schéma de modulation PAM : (a) émission (Tx) ; (b) réception (Rx).

A l'émission, le signal numérique binaire est généré par un PRBS et codé pour la correction d'erreurs en réception grâce à un codeur FEC. Ce signal est transformé en un signal numérique multiniveaux grâce au codeur PAM. Ce dernier est suréchantillonné, mis en forme et passé dans un convertisseur numérique/analogique (CNA). Ensuite, le signal filtré passe-

bas est inséré dans le modulateur (onde modulante) ; l'onde porteuse consistant en un signal optique généré par un laser optique. Le modulateur utilisé peut être un modulateur Mach-Zehnder aussi bien qu'un modulateur électro-absorbant.

A la réception, la procédure inverse est mise en place. Le signal optique est photodétecté par une photodiode (détection directe), le signal électrique résultant est filtré par un filtre passe-bas, passé dans un convertisseur analogique/numérique (CAN) et souséchantillonné. Ensuite, un décodeur PAM transforme le signal numérique multiniveaux en un signal numérique binaire et le décodage FEC est appliqué.

Ce format de modulation permet d'augmenter l'efficacité spectrale selon le nombre M de niveaux utilisés. Le premier lobe du spectre optique du signal PAM est égal à $2 \times \text{débit} / \log_2(M)$. Sur la Figure 2.10 nous présentons plusieurs cas de modulation PAM. Les lignes représentent M : PAM 2 (a, b, c) ; PAM 4 (d, e, f) ; et, PAM 8 (g, h, i). Les colonnes représentent : les constellations (a, d, g) ; le spectre optique (b, e, h) ; et, le diagramme de l'œil (c, f, i). Sur ces figures nous considérons des amplitudes réelles.

2.3 Architectures d'interconnexion intra-DC

2.3.1 Centre de données actuels

Le réseau d'interconnexion des DC actuels est basé sur la commutation électronique de paquets (EPS) (23% de la consommation d'énergie total du DC [1]). Cette architecture d'interconnexion, dite « classique », est basée sur la topologie en arbre présentée sur la Figure 2.11. Cette topologie distingue trois couches de commutation : ToR (*Top of the Rack*), agrégation et cœur ; chaque couche étant basée sur l'EPS. Les serveurs hébergés dans un rack sont connectés les uns aux autres grâce à un commutateur ToR, des connexions à 1 Gbit/s sont prévues à cet effet. Les commutateurs ToR servent en même temps à connecter la couche ToR à la couche d'agrégation, cette fois-ci avec des liens à 10 Gbit/s. Ensuite, la couche d'agrégation est connectée à la couche cœur via les commutateurs d'agrégation, toujours avec des connexions à 10 Gbit/s ; les commutateurs cœur vont assurer les communications vers l'extérieur du DC (internet) à l'aide d'un commutateur passerelle (ou *gateway*).

Afin de comparer l'architecture d'interconnexion EPS classique avec des architectures d'interconnexion optiques (que nous verrons par la suite), nous considérons n serveurs par rack, N commutateurs ToR (de $k + n$ ports), k commutateurs d'agrégation (de $c + N$ ports) et c commutateurs cœur. Nous considérons de même un total de $(k \times c) + (N \times k)$ connexions entre les différentes baies de 10 Gbit/s en fibre optique et $n \times N$ connexions à l'intérieur de chaque baie de 1 Gbit/s en câble en cuivre.

Le principal inconvénient de l'architecture d'interconnexion classique est sa consommation d'énergie. Ceci, en partie, est dû au fait que les commutateurs électriques utilisés dans les couches ToR, agrégation et cœur doivent effectuer de nombreuses conversions entre les domaines électrique et optique (et vice-versa). Par exemple, une matrice de commutation

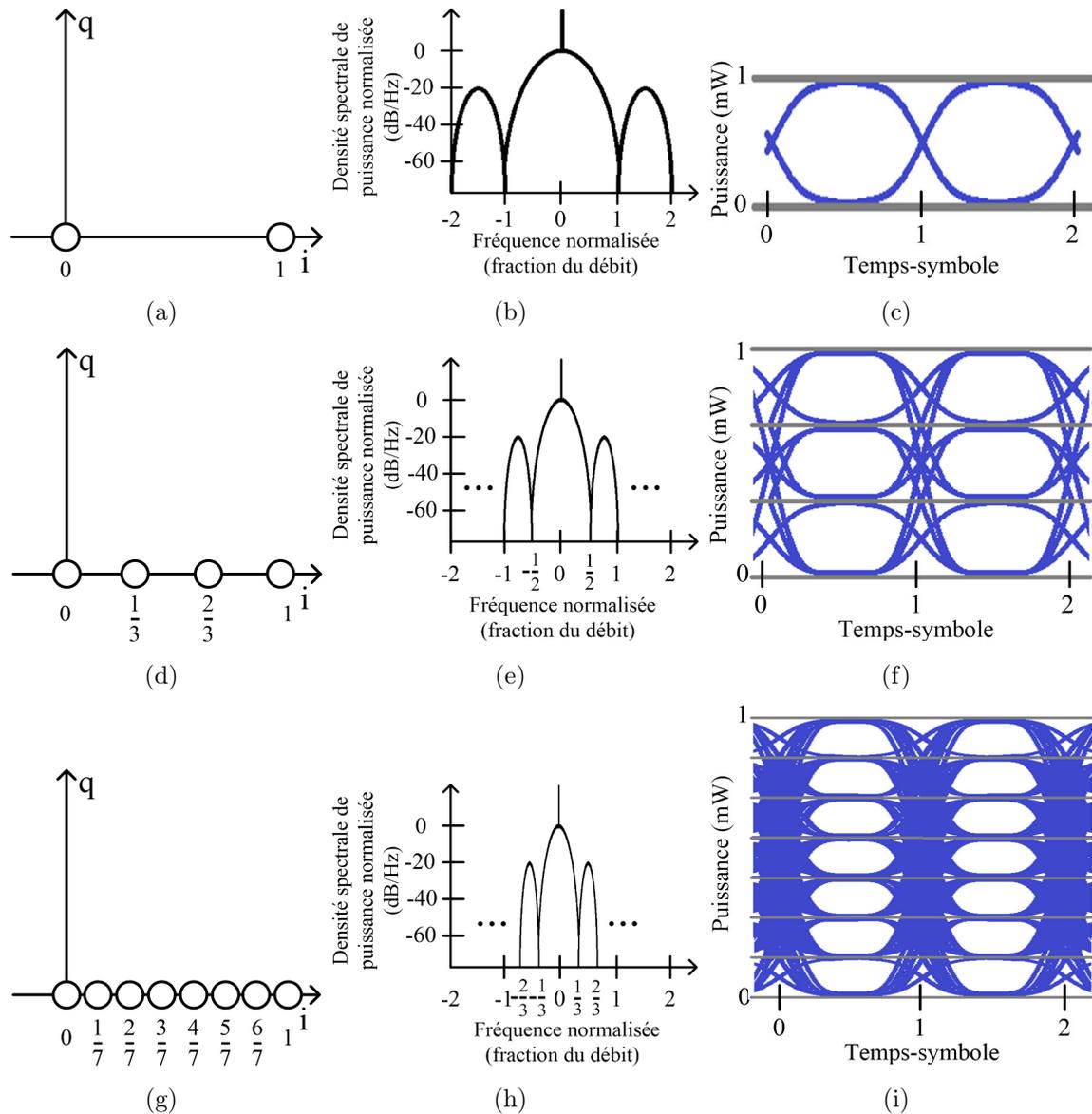


Figure 2.10. Plusieurs cas de modulation PAM. Les files représentent le niveau de modulation M : PAM 2 (a, b, c) ; PAM 4 (d, e, f) ; et, PAM 8 (g, h, i). Les colonnes représentent les constellations (a, d, g) ; le spectre optique (b, e, h) ; et, le diagramme de l'œil (c, f, i).

électrique consomme environ 25 W par port. Ainsi, la consommation d'énergie des DC actuels a un impact majeur sur l'environnement. Déjà en 2007, les DC représentaient 14% des émissions totales de gaz à effet serre (GES) des technologies de l'information (IT), équivalent à 2% des émissions GES globales. Selon les estimations réalisées par le groupe GeSI³ ces émissions devraient atteindre 18% des IT en 2020 [69].

Nous présentons sur la Figure 2.12 l'évolution du trafic dans les DC à partir de 1990. Sur cette figure nous pouvons apprécier plusieurs générations depuis l'apparition d'Internet, parmi lesquelles, les moteurs de recherche, les partages de données, les plateformes com-

3. The Climate Group on behalf of the Global eSustainability Initiative

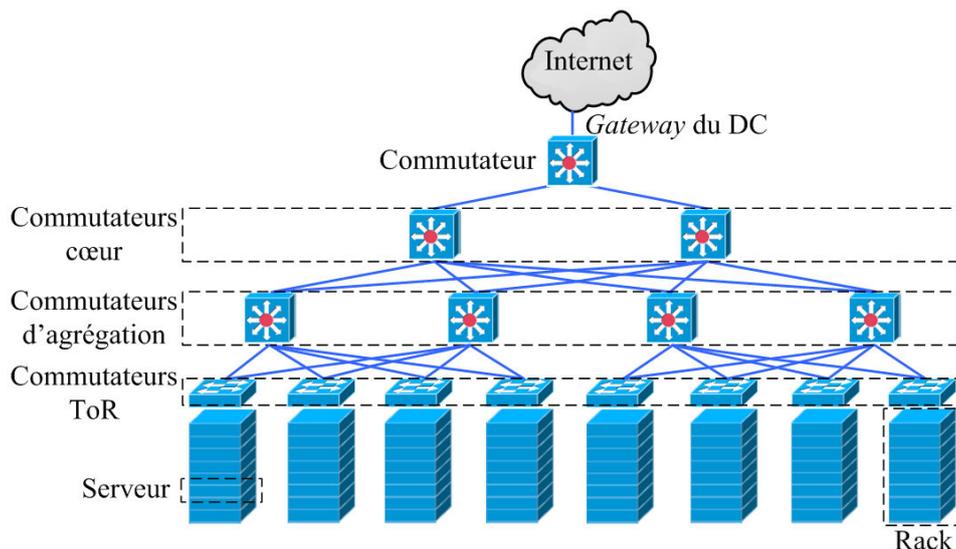


Figure 2.11. Architecture d'interconnexion EPS classique.

merciales, réseaux sociaux, etc.. De plus, la distribution de ce trafic en fonction du type de connexion est présentée sur les figures 2.13(a) et (b). La partie la plus importante de ce trafic concerne les communications internes au DC et celle-ci a une tendance à s'accroître. Par exemple, en 2016 cette dernière représentait 75%, une valeur de 77% est annoncée pour 2020.

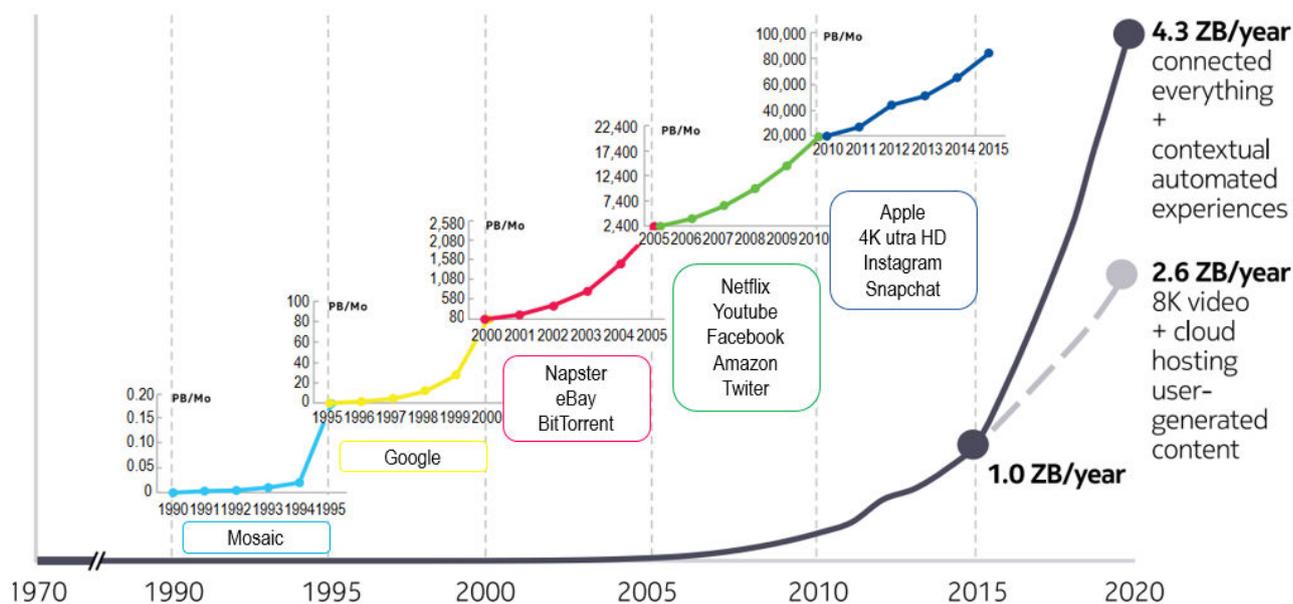
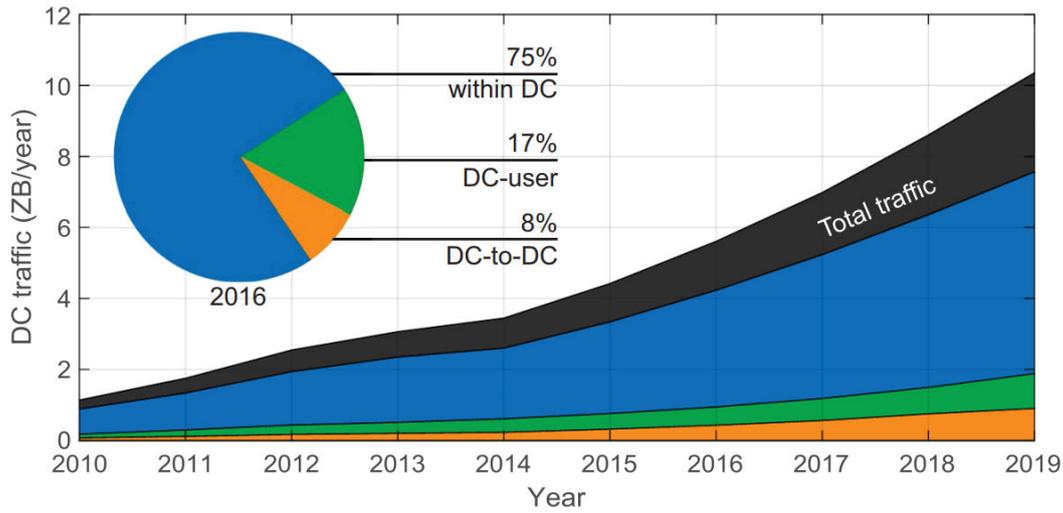
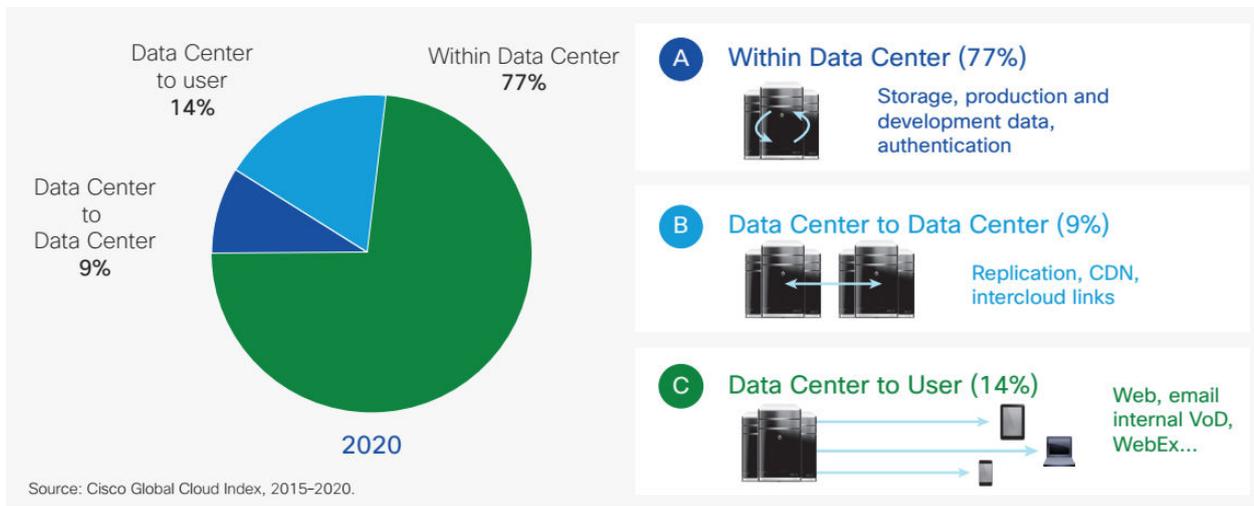


Figure 2.12. Evolution du trafic dans les DC depuis l'année 1990, trafic exprimé en zéta-octet par an [70].

Les applications émergentes, comme le *cloud computing* entre autres, requièrent des DC plus en plus puissants [73]. C'est-à-dire, que non seulement la performance des processeurs



(a)



(b)

Figure 2.13. Tendances de distribution du trafic dans le DC, (a) publiées en 2013 [71], (b) publiée en 2016 [72].

doit être améliorée, mais la capacité des réseaux doit aussi être accrue (en termes de débit total supporté par le DC) tout en diminuant leur consommation d'énergie. Le Tableau 2.4 [74]-[75] présente des tendances de la performance des DC à partir de l'année 2012. A noter, que ces tendances, qui ont été publiées en 2010, sont en accord avec celles de [70] présentées en 2016 en ce qui concerne le trafic total du DC.

Ce tableau met en évidence la nécessité d'augmenter tous les quatre ans la performance des processeurs et le débit des DC dans des facteurs de 20 et 10 respectivement juste en doublant leur consommation d'énergie. L'introduction des communications par fibre optique semble impérative, car l'optique permet d'atteindre des débits très élevés en réduisant la consommation d'énergie par rapport aux solutions actuelles [3].

Année	Performance des processeurs	Débit total du DC	Consommation d'énergie maximale
2012	10 (PFlops)	1 (Po/s)	5 (MW)
2016	100 (PFlops)	20 (Po/s)	10 (MW)
2020	1000 (PFlops)	400 (Po/s)	20 (MW)

Tableau 2.4. Tendances de la performance des processeurs, le débit total du DC et la consommation d'énergie minimale des DC à partir de l'année 2012.

2.3.2 Architectures d'interconnexion optiques

La plupart des architectures proposées ont pour but d'introduire l'optique dans les connexions entre les différentes baies (inter-racks) [76]-[82]. Cependant, d'autres architectures proposent des solutions optiques à l'intérieur de la baie (intra-rack) [83]-[84]. Des architectures encore plus récentes, proposent des solutions entre et à l'intérieur des baies [85]-[86].

En ce qui concerne les architectures inter-racks, certains auteurs proposent de faire une migration directe vers les architectures tout-optiques. La plupart étant basée sur la commutation de circuits optiques, car la capacité d'OCS peut atteindre plusieurs centaines de ports en utilisant des commutateurs à base de MEMS 3D [87]. Dans un premier temps nous étudions l'architecture d'interconnexion OCS « classique », puis Proteus [76], architecture OCS qui réduit la consommation d'énergie en regroupant les commutateurs ToR, de façon à réduire le nombre de connexions directes à la matrice de commutations basée sur des MEMS.

Des variantes d'OCS utilisent des transmissions flexibles en termes de largeur de bande ou formats de modulation [66],[85]. Ceci permet d'augmenter ou de baisser le débit selon besoin et ainsi d'économiser de l'énergie sous certaines conditions. Nous présentons l'architecture EODCN (*Energy-Efficient-Elastic-Optical-Interconnect Architecture for Data Centers*), avec une structure commune tout-optique pour assurer les connexions intra et inter-racks, avec un plan de contrôle basé sur une structure optique indépendante de celle des données.

Nonobstant, les approches basées sur l'OCS répondent mal aux besoins d'un trafic variable dans le temps. Néanmoins, dû à son potentiel d'interconnexion, OCS est considéré dans des solutions d'interconnexion hybrides (électrique-optique). OCS serait utilisé pour le transfert de longues chaînes de données entre les commutateurs ToR ; tandis que l'EPS, pour le trafic de fond. Nous étudions HELIOS [77], architecture d'interconnexion (inter-rack) hybride qui considère un système de contrôle commun pour basculer entre l'EPS et l'OCS [88].

Ensuite, nous considérons des architectures d'interconnexions inter-racks basées sur la commutation de paquets optiques. Cette technologie ayant comme principale difficulté la fabrication des dispositifs tout-optiques qui puissent faire la commutation de paquets en des temps raisonnables, de l'ordre de la nanoseconde, [89]-[90]. Toutefois, plusieurs techniques d'OPS, basées sur différents composants optiques, ont été récemment présentées et méritent d'être étudiées au cours de cette thèse : DOS [78], Petabit [79], IRIS [80] et *Broadcast and*

Select [81].

Des architectures d'interconnexion basées sur des émetteurs optiques rapidement accordables en longueur d'onde ont été récemment étudiées, en particulier pour des applications dans le domaine des réseaux métropolitains. Ces techniques incluent POADM (*Packet Optical Add and Drop Multiplexer*) [91], proposée par Alcatel-Lucent, et TWIN (*Time-domain Wavelength Interleaved Networks*), proposée à l'origine par Lucent [92].

L'application de POADM pour des connexions intra DC a été récemment proposée sous le nom d'OSS (*Optical Slot Switching*) [82]. OSS offre uniquement des connexions inter-racks tout-optiques. POADM exige également des portes optiques rapides, mais qui sont toujours moins matures que les sources accordables. En outre, cette technologie utilise une topologie en anneau, et les paquets doivent attendre en utilisant une FDL (*Fiber-Delay Line*) dans chaque nœud pour le traitement du paquet de contrôle. Cela génère un temps de latence, qui peut être négligé dans les applications de télécommunications, mais pas pour les applications courte distance, comme les communications intra-DC.

Dans TWIN une longueur d'onde en particulier est attribuée à chaque nœud de destination, afin de recevoir des paquets uniquement sur cette longueur d'onde. Chaque destination est ainsi à la racine d'un arbre coloré. La Figure 2.14 montre un exemple de réseau TWIN. Les nœuds intermédiaires sont purement passifs par rapport à l'aiguillage des paquets, ils peuvent être configurés pour créer l'ensemble des arbres colorés, par exemple en utilisant des WSS. La complexité du réseau de paquets est localisée dans les nœuds source et de destination. Un plan de contrôle est nécessaire pour gérer les contraintes de collision dans le réseau et gérer les priorités à l'émission. Par exemple, dans la Figure 2.14 les paquets sont émis dans l'ordre suivant : de S1 vers D1 ; de S2 vers D1 ; de S3 vers D2 ; de S2 vers D3 ; de S3 vers D3. Des travaux ont montré qu'un contrôle centralisé est plus efficace qu'un contrôle distribué [93]. Dans le cas où la topologie comporte un point central le contrôle est proche de celui d'un réseau d'accès de type PON.

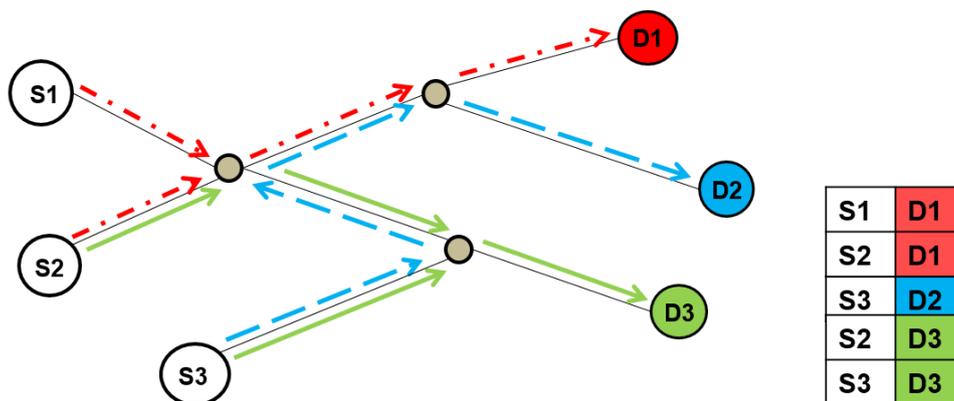


Figure 2.14. Architecture d'interconnexion TWIN.

POITR (*Passive-Optical-Pod Interconnect at Top of the Rack*) [83] est une architecture

d'interconnexion pour DC purement passive dérivée de TWIN. Cette architecture d'interconnexion offre des connexions tout-optiques à l'intérieur de la baie. Ses auteurs proposent de combiner POITR avec Proteus, afin d'assurer des solutions tout-optiques dans la globalité du DC.

Une autre option dérivée de TWIN, qui dépend aussi d'un réseau purement passif, a été présentée dans [86] sous le nom de POPI (*Passive Optical Pod Interconnect*). Cette architecture d'interconnexion n'utilise pas de commutateur ToR et relie tout les serveurs du DC à l'aide d'un coupleur passif. POPI est donc un cas de réseau TWIN avec un point central. Par ailleurs, elle ne fait pas de distinction entre connexions intra-ou inter-racks.

2.3.2.1 Architectures d'interconnexion inter-racks

2.3.2.1.1 Architecture d'interconnexion OCS classique

Cette architecture remplace les commutateurs de paquets électriques des couches agrégation et cœur de l'architecture EPS classique par des commutateurs de circuits optiques (représentés en couleur verte sur la Figure 2.15). Grâce au multiplexage en longueur d'onde, les liaisons inter-racks mono-canal sur fibre optique à 10 Gbit/s vont augmenter leur capacité : plusieurs émetteurs et récepteurs (plusieurs canaux WDM) vont être mis en place au niveau des commutateurs ToR. La Figure 2.15(a) présente le schéma d'interconnexion de cette architecture. La Figure 2.15(b) montre les connexions entre les commutateurs ToR et d'agrégation à l'aide des (dé)multiplexeurs WDM. Les figures 2.15(c) et (d), aident à expliquer le fonctionnement des commutateurs OCS utilisant des MEMS. La Figure 2.15(c) montre la connexion entre l'entrée 1 et la sortie 1. Le faisceau lumineux de l'entrée 1 passe par une lentille et il est réfléchi plusieurs fois par différents miroirs : 1. miroir mobile 1, 2. miroir fixe et 3. miroir mobile 3, qui redirige la lumière vers la sortie 1. La Figure 2.15(d) montre la connexion entre l'entrée 1 et la sortie 2. Suite à la rotation du miroir mobile 1, le faisceau lumineux est réfléchi par : 1. miroir mobile 1 qui a changé de position, 2. miroir fixe et 3. miroir mobile 2, qui redirige la lumière vers la sortie 2.

Cette architecture utilise N commutateurs ToR, k commutateurs d'agrégation et c commutateurs cœur, de $k+n$, $N+c$ et $c+1$ ports chacun, respectivement. Ayant $w \times (N \times k + c)$ connexions en fibre optique à 10 Gbit/s chacune, avec w longueurs d'onde en multiplexage WDM.

2.3.2.1.2 Proteus [76] par A. Singla et al., 2010

Dans cette architecture d'interconnexion chaque commutateur ToR, ayant plusieurs transpondeurs (un par serveur), est connecté à la matrice de commutation basée sur des MEMS. Les longueurs d'onde en provenance des commutateurs ToR sont multiplexées et routées vers un WSS, puis elles sont envoyées vers quatre différents groupes de sortie. Chaque groupe est connecté à un port différent de la matrice de commutation à l'aide des circulateurs optiques (voir le schéma d'interconnexion de Proteus sur la Figure 2.16). De cette façon, plusieurs

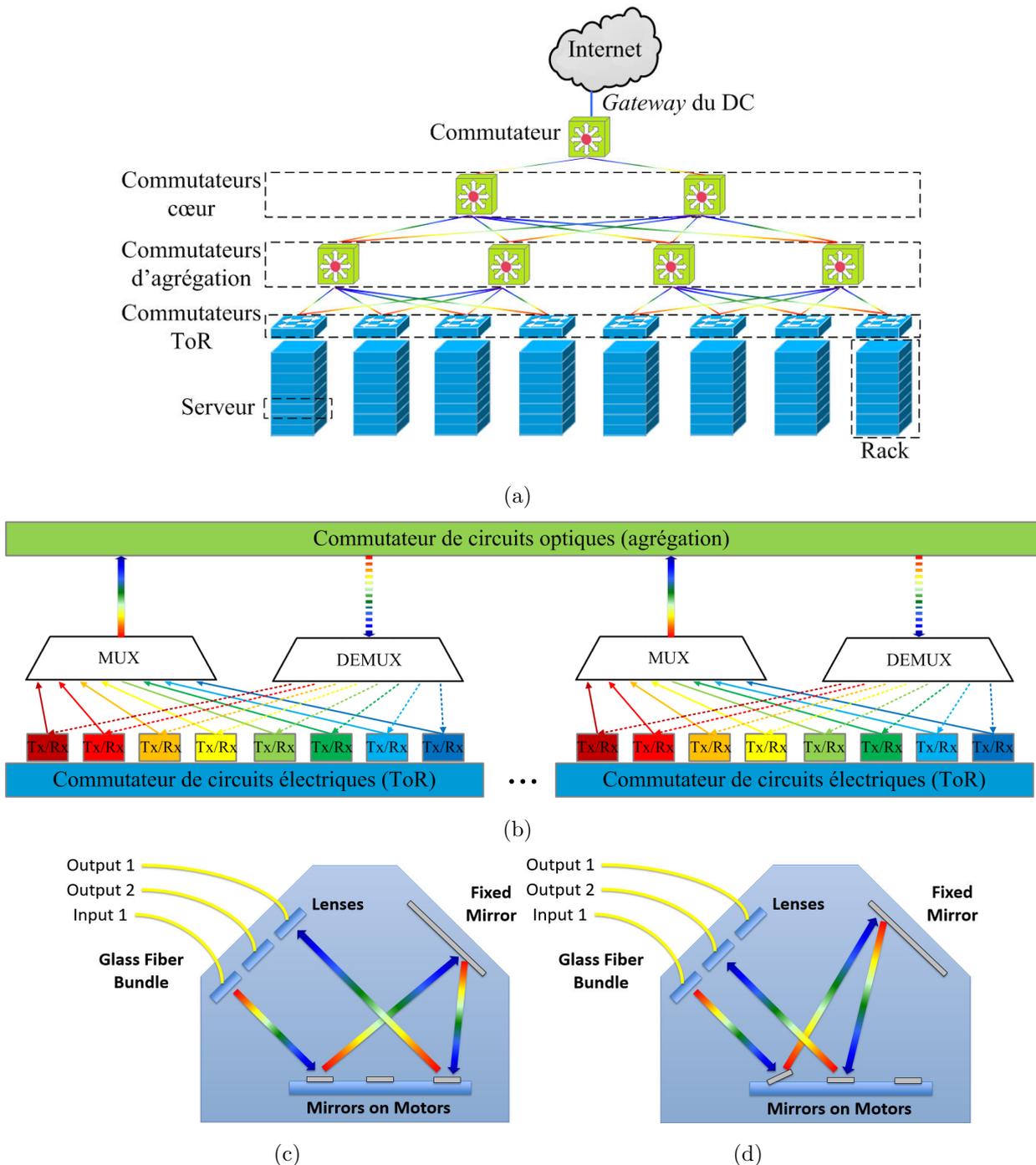


Figure 2.15. Architecture d'interconnexion inter-racks basée sur OCS classique : (a) schéma d'interconnexion, commutateurs OCS représentés en vert ; (b) connexions entre les couches d'agrégation et ToR ; (c) et (d) fonctionnement de l'OCS en utilisant des MEMS [94].

connexions point-à-point entre différents commutateurs ToR sont établies. Si jamais il y a une demande de connexion entre 2 commutateurs ToR qui n'ont pas une connexion directe, la communication se fait *hop-by-hop*. Pour la réception, les longueurs d'onde suivent le processus inverse.

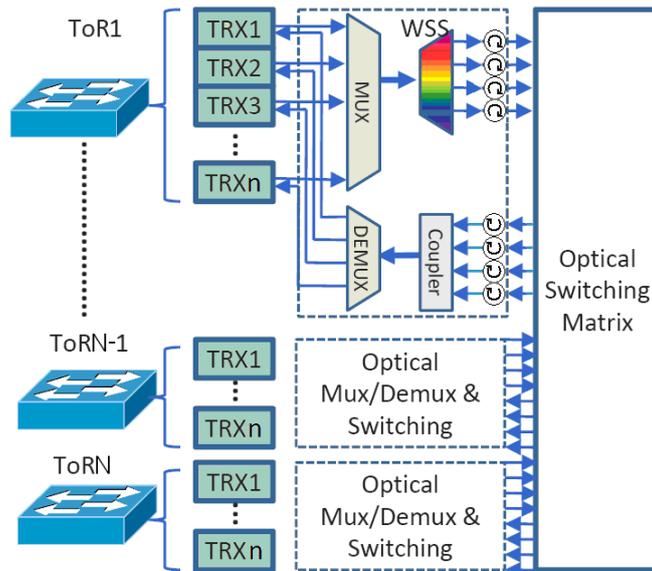


Figure 2.16. Architecture d'interconnexion Proteus.

Une des principales exigences pour assurer le bon fonctionnement du réseau Proteus est de trouver la configuration optimale de la matrice de commutation à base de MEMS pour chaque modèle de trafic. Les auteurs de [76] proposent de faire ceci en utilisant la programmation linéaire entière.

Proteus utilise N commutateurs ToR de $n + n$ ports chacun ; $2 \times N$ (dé)multiplexeurs ($1 : n/n : 1$) ; N WSS de 4 ports de sortie ; $2 \times N \times 4$ circulateurs, N coupleurs ($4 : 1$) et une matrice de MEMS de $2 \times 4 \times N$ ports.

2.3.2.1.3 HELIOS : Hybrid-Electrical/Optical Switch [77] par N. Farrington et al., 2010

HELIOS est une architecture inter-rack hybride EPS/OCS. Son schéma d'interconnexion est présenté sur la Figure 2.17. Cette architecture d'interconnexion ne comporte pas d'agrégation et connecte les commutateurs ToR (EPS) directement aux commutateurs cœur (soit EPS ou OCS). Un contrôleur SDN commun permet de basculer entre les commutateurs EPS et OCS [88]. De plus, les auteurs de HELIOS ont mené des études sur l'optimisation du rapport entre les commutateurs EPS/OCS, en considérant plusieurs canaux WDM, en termes de coût et de consommation d'énergie. Les résultats de leurs études ont montré que le meilleur compromis coût/consommation d'énergie est possible en utilisant 10% d'EPS et 90% d'OCS, pour 4 et 8 canaux WDM.

HELIOS utilise N commutateurs ToR de $c + n$ ports chacun ; $0, 1 \times c$ et $(0, 9 \times c)/8$ (pour 8 longueurs d'ondes différentes) commutateurs cœur EPS et OCS, respectivement, de $N + 1$ ports chacun.

2.3.2.1.4 DOS : Data-Center-Optical Switch [78] par X. Ye et al., 2010

un seul commutateur optique (pas de retard dû à la mémoire tampon des commutateurs électriques). Ses principaux inconvénients sont la consommation d'énergie très élevée de la mémoire tampon électrique utilisée en cas de collision et la complexité de fabrication des convertisseurs de longueur d'onde reconfigurable.

DOS utilise N commutateurs ToR; N LE; N convertisseurs de longueur d'onde; un AWG de $N + 1/N + 1$ entrées/sorties; un plan de commande et une mémoire tampon de a convertisseurs O/E-E/O (a étant la probabilité de collision).

2.3.2.1.5 Petabit [79] par K. Xia et al., 2010

Petabit, réseau d'interconnexion pour DC, est basé sur le projet PetaStar présenté en 2003 dans [95]. Cette architecture d'interconnexion est également basée sur des AWG et des convertisseurs de longueur d'onde. Son schéma d'interconnexion est présenté sur la Figure 2.19.

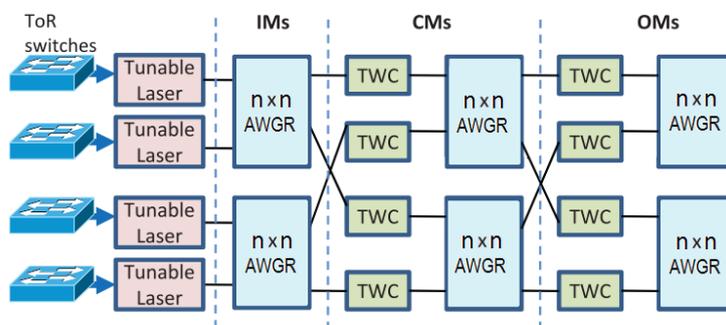


Figure 2.19. Architecture d'interconnexion inter-baies Petabit [73]. *IMs* : *Input Modules*, *CMs* : *Central Modules*, *OMs* : *Output Modules*, *TWC* : *Tunable-Wavelength Converter*, *AWGR* : *Arrayed-Waveguide-Grating Routing*.

Cette architecture d'interconnexion est organisée en trois étages de commutation (réseau de Clos, 1953). La commutation optique est combinée avec de l'ordonnancement et de la mise en mémoire tampon électriques. Avant la transmission, les paquets sont stockés dans la mémoire tampon et ensuite classés et ordonnés par rapport à leur information de destination dans une file d'attente de sortie virtuelle. Un algorithme d'ordonnancement détermine les ports à utiliser à chaque étage de commutation. Petabit présente une faible latence de l'ordre de 200 ns.

Petabit utilise N commutateurs ToR; $(3 \times N)/n$ AWG de n/n entrées/sorties et $2 \times N$ convertisseurs de longueur d'onde.

2.3.2.1.6 IRIS : Integrated Router Interconnected Spectrally [80] par J. Gripp, 2010

IRIS est aussi une architecture d'interconnexion basée sur des AWG et des convertisseurs de longueur d'onde. Son schéma d'interconnexion est présenté sur la Figure 2.20.

Elle compte aussi trois étages de commutation optique; les premier et troisième étages

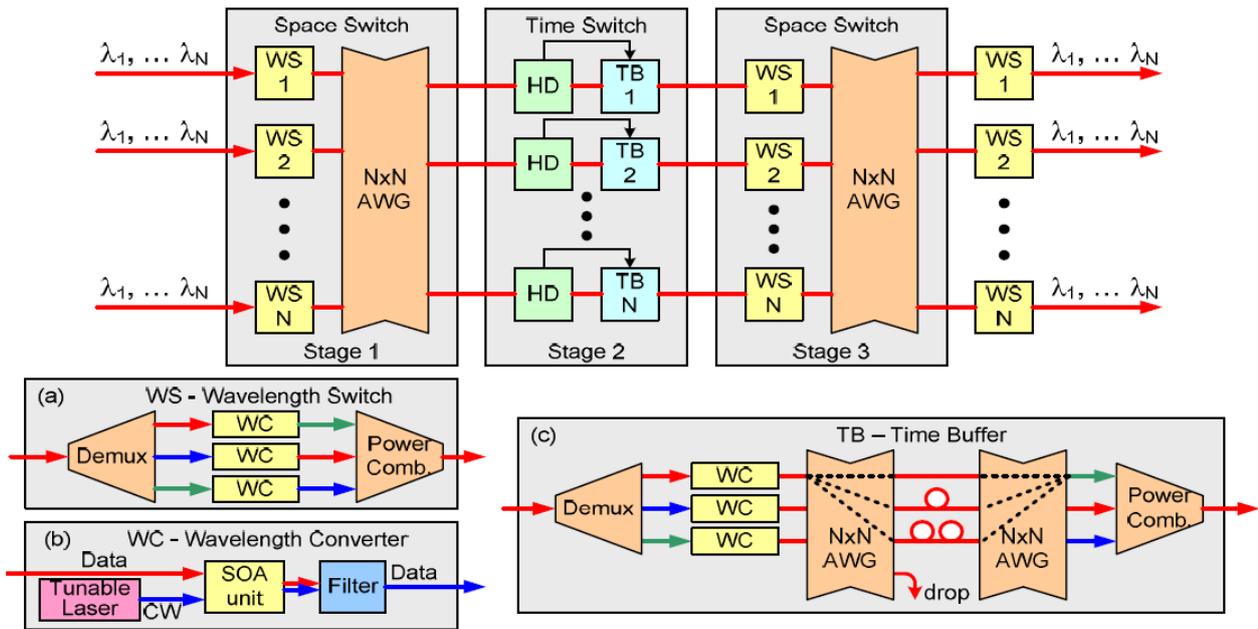


Figure 2.20. Architecture d'interconnexion inter-baies IRIS [80] : (a) commutateur de longueur d'onde (*WS* : *Wavelength Switch*) ; (b) détecteur de l'entête (*HD* : *Header Detector*) ; (c) module de retard (*TB* : *Time Buffer*).

sont spatiaux et, le deuxième, temporel. Le premier étage consiste en la connexion des commutateurs ToR vers des commutateurs de longueurs d'onde. Les WS sont des dispositifs créés à partir d'une paire de (dé)multiplexeurs (1 : 3/3 : 1) et de trois convertisseurs de longueur d'onde, selon la Figure 2.20(a). Les sorties du commutateur de longueurs d'ondes sont connectées à un AWG qui sera ensuite connecté au deuxième étage de commutation. L'entête du paquet sera détectée et le paquet sera envoyé dans un tampon optique qui retardera le passage vers l'étage suivant si nécessaire.

IRIS utilise N commutateurs ToR ; $6 \times N$ (dé)multiplexeurs ; 4 AWG de N/N entrées/sorties ; $9 \times N$ convertisseurs de longueur d'onde et N HD.

2.3.2.1.7 Broadcast and Select [81] par R. Luijten, 2004

L'architecture d'interconnexion *Broadcast and Select* a été conçue pour la connexion entre 64 commutateurs ToR. Le schéma d'interconnexion est présenté sur la Figure 2.21. Cette architecture d'interconnexion comporte deux étages : diffusion (*broadcast*) et sélection (*select*). Des modules de mémoire tampon à l'entrée ; 64, à la sortie⁴ du réseau, ainsi qu'un canal de contrôle par module, sont mis en place afin éviter les collisions.

Dans un premier temps, les commutateurs ToR sont regroupés à l'aide de 8 multiplexeurs 8 : 1 dans 8 fibres optiques, chaque commutateur ToR ayant une longueur d'onde assignée parmi les huit possibilités. Ensuite, ces longueurs d'onde sont amplifiées (grâce à un amplificateur à fibre dopée à l'erbium), diffusées (grâce à un coupleur 1 : 128) et connectées aux

4. Avec deux récepteurs pour redondance.

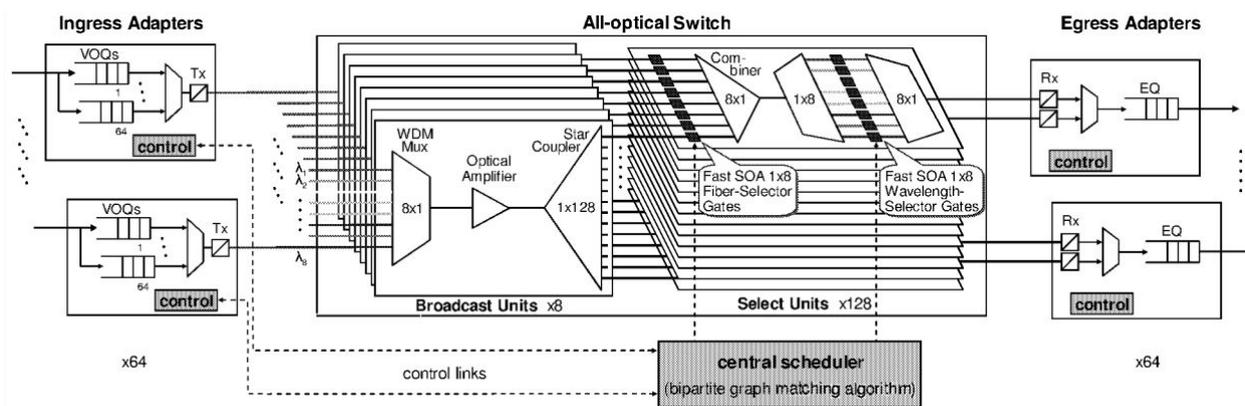


Figure 2.21. Architecture d'interconnexion *Broadcast and Select* [81].

modules de sélection.

Dans un deuxième temps, les sorties de chaque module de diffusion (8 modules) sont regroupées grâce à un coupleur 1 : 8 (sorties auparavant passées par des portes de sélection basées sur des SOA ; à noter que parmi les 2048 SOA de cette architecture, seulement 128 sont actifs). Puis, la sortie sélectionnée est dé-multiplexée et un deuxième passage à travers les portes de sélection est prévu, suivi par un multiplexage 8 : 1.

2.3.2.1.8 OSS : Optical-Selective Switch [82] par M. A. Mestre, 2014

Cette architecture d'interconnexion inter-racks est basée sur POADM [91]. Elle consiste en un réseau anneau sous lequel les commutateurs ToR sont connectés grâce à un nœud OSS. Le réseau OSS et le schéma du nœud OSS sont présentés sur les figures 2.22(a) et 2.22(b), respectivement.

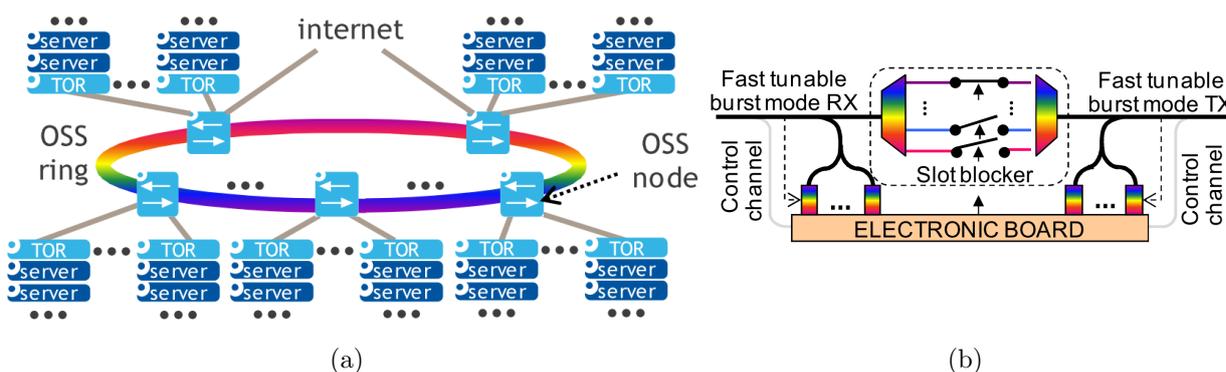


Figure 2.22. (a) Architecture d'interconnexion basée sur OSS ; (b) nœud OSS [82].

Chaque nœud OSS est composé d'un émetteur rapidement accordable en longueur d'onde et d'un récepteur avec détection cohérente⁵ ; d'une paire de (dé)multiplexeurs, entre lesquelles plusieurs (selon le nombre de longueurs d'onde) portes logiques à base de SOA

5. Nous reviendrons sur le mode de détection dans la section destinée aux formats de modulation.

bloquent (ou pas) certains paquets afin d'éviter les collisions ; deux coupleurs pour l'insertion/extraction de paquets ; et un canal de contrôle. Chaque longueur d'onde est partagée par deux nœuds OSS, permettant un maximum de 80 nœuds OSS connectés à l'anneau.

2.3.2.2 Architecture d'interconnexion intra-rack

2.3.2.2.1 POITR : Passive-Optical-Pod Interconnect at Top of the Rack [83] par Y. Gong, 2014

Cette architecture d'interconnexion est basée sur TWIN [92]. Il s'agit d'un réseau optique purement passif, sa consommation d'énergie est donc très faible. Les auteurs proposent deux options d'interconnexion, présentées sur la Figure 2.23.

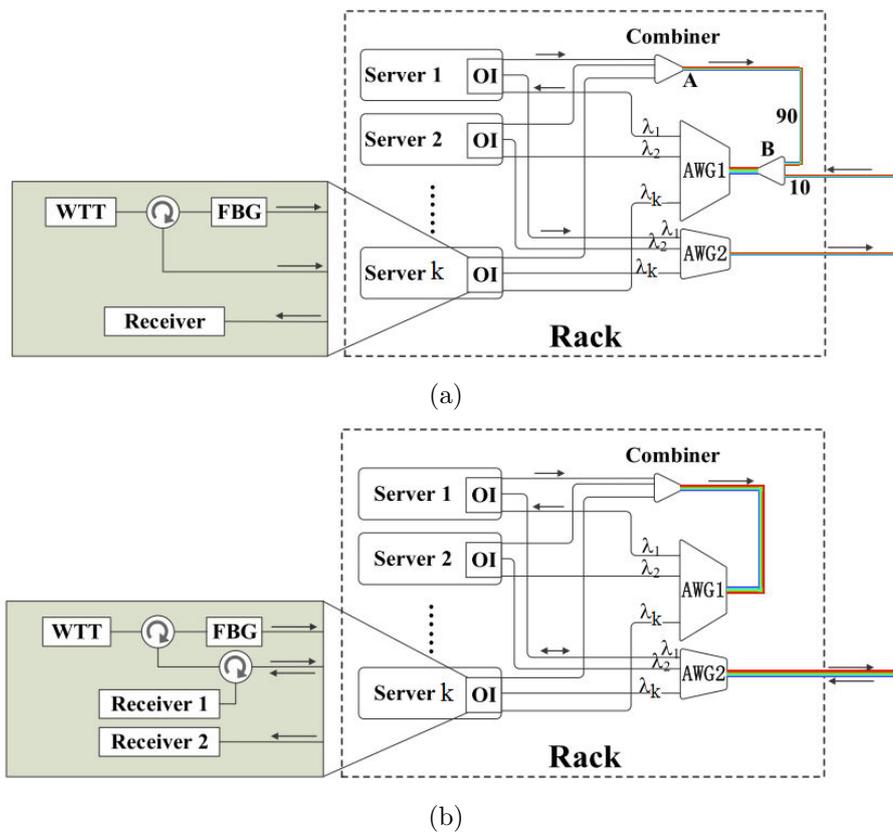


Figure 2.23. Architecture d'interconnexion POITR [83] : (a) 1 Rx ; (b) 2 Rx. *OI* : *Optical Interface*, *WTT* : *Wavelength-Tunable Transmitter*

La première option (Figure 2.23(a)) utilise dans chaque serveur i une interface optique constituée par un émetteur rapidement accordable, un circulateur, un filtre de Bragg centré sur la longueur d'onde i et un récepteur. Ce réseau utilise deux coupleurs (A et B) et deux (dé)multiplexeurs (AWG1 et AWG2).

Pour les connexions intra-racks chaque serveur doit accorder son émetteur à une longueur d'onde différente de celle pour laquelle son récepteur a été ajusté, de telle sorte de ne pas être réfléchi par le filtre de Bragg. Le coupleur A combine tous les flux venant des serveurs

du même rack. Ensuite, le coupleur B combine ce flux avec les flux venant des autres racks. Chacune des sorties du AWG1 est connectée à un serveur en accord avec la longueur d'onde de son récepteur.

D'autre part, pour les connexions inter-racks, les serveurs doivent accorder leur émetteur à la longueur d'onde pour laquelle leur propre récepteur a été ajusté. De cette façon, elle sera réfléchiée par le filtre de Bragg et, grâce au circulateur, passera directement vers l'AWG2 qui va adresser ce flux à l'extérieur du rack. Les auteurs utilisent Proteus [76] pour adresser ce flux vers le rack de destination et à nouveau POITR pour l'adresser au serveur de destination une fois trouvé le rack de destination.

La deuxième option (Figure 2.23(b)), supprime le coupleur B et ajoute un deuxième circulateur et un deuxième récepteur au niveau de l'OI. De cette façon, les pertes d'insertion dues aux coupleurs diminuent considérablement. Maintenant l'AWG1 permet d'adresser uniquement les connexions intra-racks ; et l'AWG2, uniquement les connexions inter-racks.

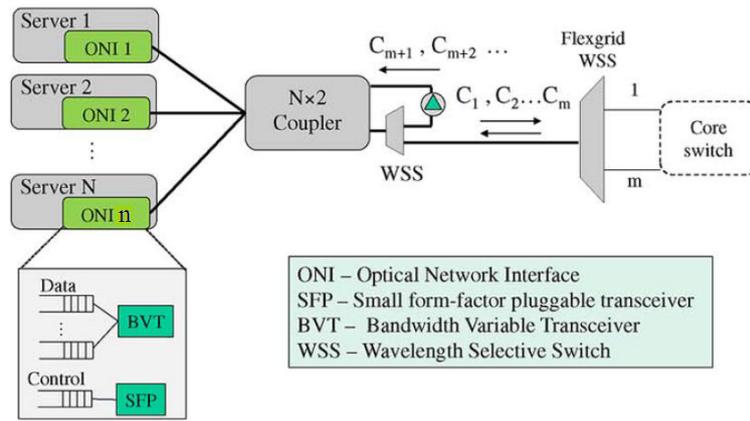
Une autre architecture d'interconnexion en parallèle est considérée pour transmettre les informations utiles au plan de commande de POITR. Il s'agit de la structure de contrôle utilisée par EODCN [85], nous la décrivons dans la section suivante.

L'ensemble POITR + Proteus utilise $n \times N$ interfaces optiques ; $2 \times N$ (dé)multiplexeurs ; N ou $2 \times N$ coupleurs (option a ou b) ; N commutateurs ToR de $n + n$ ports chacun ; $2 \times N$ (dé)multiplexeurs ($n : 1/1 : n$) ; N WSS de 4 ports de sortie ; $2 \times N \times 4$ circulateurs, N coupleurs ($4 : 1$) et une matrice de MEMS de $2 \times 4 \times N$ ports ; $(n + 1) \times N$ transpondeurs SFP et N plans de commande locaux et 1 global.

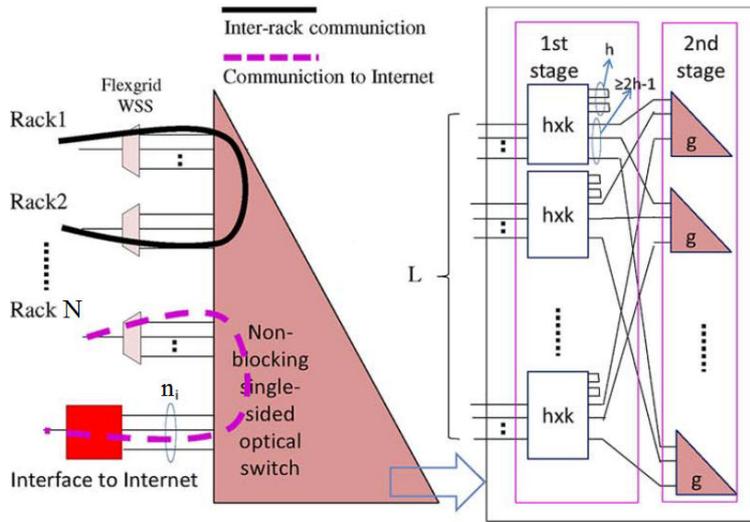
2.3.2.3 Architectures d'interconnexion inter et intra-racks

2.3.2.3.1 EODCN : Elastic-Optical-DC Network [85] par M. Fiorani et al., 2014

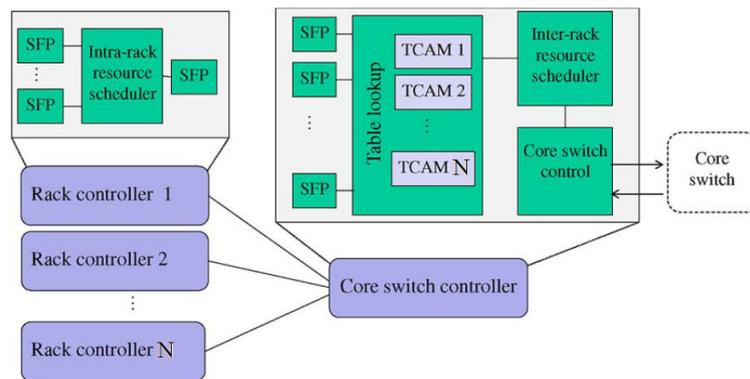
EODCN est une architecture d'interconnexion basée sur l'OCS. La Figure 2.24(a) présente la structure du réseau intra-rack. Chaque serveur i est connecté aux réseaux de données et de contrôle grâce à une interface optique (ONI) constituée d'un transpondeur de largeur de bande variable (BVT) pour la transmission de données et d'un transpondeur SFP pour le contrôle. Les BVTs sont connectés à un coupleur passif $n : 2$, dont une branche est connectée à un WSS de largeur de bande variable $1 : 2$ qui va rediriger le flux intra-rack vers la deuxième sortie du coupleur pour le diffuser vers d'autres serveurs du même rack. L'autre sortie du WSS est ainsi démultiplexée à l'aide d'un deuxième WSS de largeur de bande variable qui sera connecté au commutateur cœur pour les connexions inter-racks. Le commutateur cœur est constitué de deux étages de commutation basée sur des matrices de MEMS selon la Figure 2.24(b).



(a)



(b)



(c)

Figure 2.24. Architecture d'interconnexion EODCN [85] : (a) structure intra-rack ; (b) structure inter-racks ; (c) structure de contrôle. *ONI* : *Optical-Network Interface*. *BVT* : *Bandwidth-Variable Transceiver*. *TCAM* : *Ternary-Content-Addressable Memory*.

La structure de contrôle est présentée sur la Figure 2.24(c). Tous les serveurs d'un même rack sont connectés directement vers un plan de commande local. Ce dernier a de même un

émetteur SFP qui sera connecté au plan de contrôle global qui est constitué par : un module de calcul de chemins grâce à des TCAM; un module d'ordonnancement inter-racks; et un module de reconfiguration des matrices de commutation.

L'ensemble EODCN utilise $n \times N$ interfaces optiques; N coupleurs $1 : 2$; $2 \times N$ WSS de largeur de bande flexible; N isolateurs et une matrice de MEMS de $1 + L/h + 2 \times (h - 1)$ ports; N transpondeurs SFP et N plans de contrôle locaux et un global.

2.3.2.3.2 POPI : Passive-Optical Pod Interconnect [86] par R-M. Indre, 2014

POPI est une architecture d'interconnexion qui n'utilise pas de commutateurs ToR. Cette architecture d'interconnexion est également basée sur TWIN et ressemble à POITR sur certains aspects. Chaque serveur est constitué d'un émetteur rapidement accordable en longueur d'onde et d'un récepteur ajusté à une longueur d'onde fixe. Les auteurs ont prévu deux modules de contrôle (1 pour redondance) afin d'éviter les collisions de paquets et un *gateway* pour assurer les connexions hors du DC. Serveurs, contrôleurs et *gateway* sont connectés directement grâce à un coupleur passif $(N + 3) \times (N + 3)$. POPI choisit le destinataire en fonction de la longueur d'onde accordée par l'émetteur, soit un rack ou un serveur. Dans le premier cas (option 1), un rack peut partager une certaine longueur d'onde parmi tous ses serveurs et identifier le serveur destinataire en fonction de l'entête du paquet, (voir racks 1 et 2 de la Figure 2.25(a)). Le deuxième cas (option 2), réserve une longueur d'onde par serveur (voir rack N de la Figure 2.25(a)), autrement dit, POPI peut se passer de la notion de rack et ne plus faire la distinction entre connexions intra ou inter-racks.

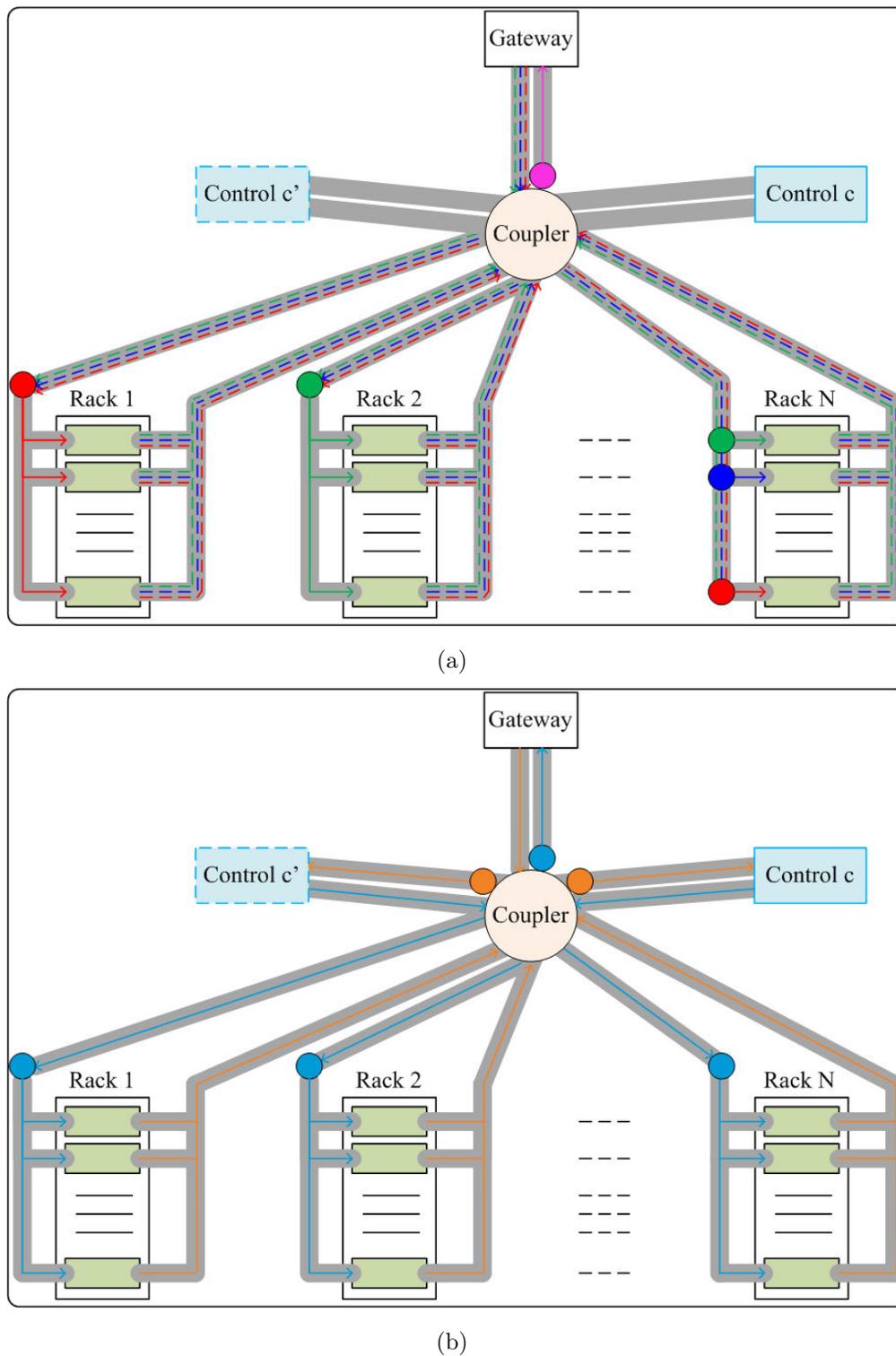


Figure 2.25. Architecture d'interconnexion POPI : (a) transmission des données ; (b) transmission des messages de contrôle.

Les serveurs demandent l'autorisation pour établir des connexions entre eux en envoyant des messages de *report* au contrôleur sous la longueur d'onde λ_r (représentée en couleur orange sur la Figure 2.25(b)). Le contrôleur de son côté, répond aux serveurs avec des

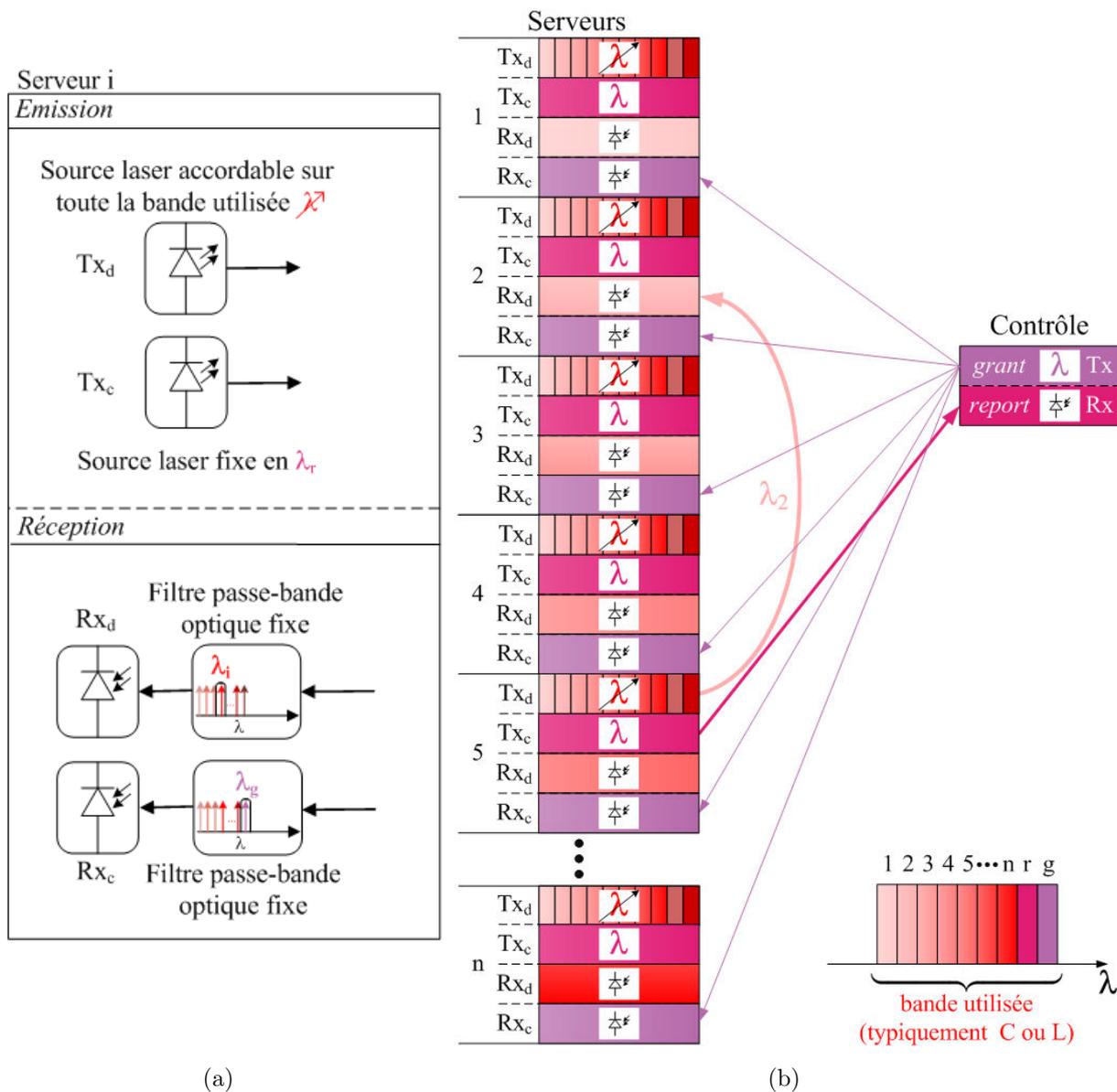


Figure 2.26. Réseau POPI : (a) interfaces de données et de contrôle; (b) exemple de connexion entre les serveurs 5 et 2, Tx_d et Tx_c représentent les émetteurs de données et de contrôle; Rx_d et Rx_c , les récepteurs de données et de contrôle, respectivement.

messages de *grant* sous la longueur d'onde λ_g (représentée en couleur bleue ciel sur la Figure 2.25(b)), autorisant la connexion et précisant le temps et la durée de la transmission.

Les auteurs de POPI ne précisent pas si la structure de contrôle est supporté par le réseau de données ou si elle est mise en place grâce à un réseau en parallèle (comme pour EODCN ou POITR). Il n'est pas précisé si les canaux de contrôle sont générés par des émetteurs spécifiques ou par les émetteurs des données. Dans un premier temps, afin de pouvoir estimer la consommation d'énergie de POPI, nous allons considérer des transpondeurs SFP au niveau de chaque serveur (en plus des émetteurs) et un réseau de données et de contrôle partagé. La

Figure 2.26(a) présente les interfaces de données et de contrôle au niveau de chaque serveur. Ainsi que la Figure 2.26(b) montre un exemple d'interconnexion entre les serveurs 5 et 2⁶.

En conséquence POPI utilise (option 1) $n \times N$ émetteurs rapidement accordables ; $n \times N$ récepteurs ; $n \times N + 3$ transpondeurs SFP ; un coupleur passif $2 \times N + 3 : 2 \times N + 3$; deux contrôleurs et un *gateway*.

2.3.3 Consommation d'énergie des différentes architectures d'interconnexion

Afin de pouvoir estimer la consommation d'énergie des différentes architectures d'interconnexion pour DC présentées précédemment, les tableaux 2.5 et 2.6 présentent respectivement les expressions à utiliser pour chaque architecture et la consommation d'énergie par composant.

Ainsi, nous considérons $N = 32$ racks qui hébergent $n = 32$ serveurs chacun, $k = 32$ et $c = 16$, commutateurs d'agrégation et cœur, pour comparer la consommation d'énergie de chaque architecture d'interconnexion. Les résultats de ce calcul sont présentés sur la Figure 2.27.

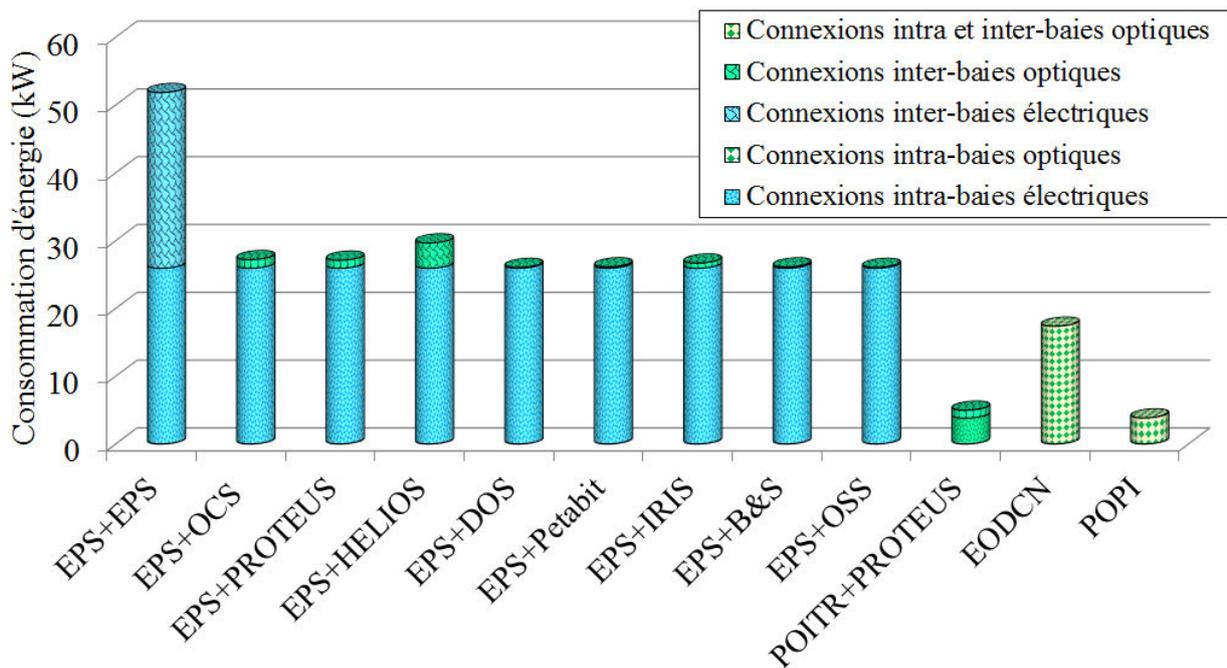


Figure 2.27. Consommation d'énergie des architectures d'interconnexion pour DC.

A partir de cette figure nous pouvons mettre en évidence que les solutions optiques peuvent réduire d'un facteur de dix la consommation d'énergie. Nous pouvons aussi constater que les architectures d'interconnexion inter-baies tout-optiques, qui ne considèrent pas de

6. Pour simplicité, nous avons omis le coupleur passif (toujours présent dans l'architecture d'interconnexion).

Architecture	Connexions intra-baie	Connexions inter-baies
EPS + EPS	$N \times [n \times P_{TRx-élec} + (k+n) \times P_{EPS}]$	$k \times [N \times P_{TRx-élec} + (N+c) \times P_{EPS}]$
EPS + OCS	$N \times [n \times P_{TRx-élec} + (k+n) \times P_{EPS}]$	$k \times [N \times P_{TRx-opt} + (N+c) \times P_{OCS}]$
EPS + PROTEUS	$N \times (n \times P_{TRx-élec} + 2n \times P_{EPS})$	$N \times (n \times P_{TRx-opt} + 4P_{OCS} + P_{WSS})$
EPS + HELIOS	$N \times [n \times P_{TRx-élec} + (k+n) \times P_{EPS}]$	$0,1 (k \times [N \times P_{TRx-élec} + (N+c) \times P_{EPS}])$ $+0,9 (k \times [N \times P_{TRx-opt} + (N+c) \times P_{OCS}])$
EPS + DOS	$N \times [n \times P_{TRx-élec} + (n+1) \times P_{EPS}]$	$N \times [P_{TRx-opt} + P_{WC}$ $+a \times (P_{O-E} + P_{E-O} + P_{SDRAM})]$
EPS + Petabit	$N \times [n \times P_{TRx-élec} + (n+1) \times P_{EPS}]$	$N \times (P_{Tx-accordable} + 2P_{WC})$
EPS + IRIS	$N \times [n \times P_{TRx-élec} + (n+1) \times P_{EPS}]$	$N \times (P_{TRx-opt} + 3P_{WC})$
EPS + B&S	$N \times [n \times P_{TRx-élec} + (n+1) \times P_{EPS}]$	$N \times (P_{TRx-opt} + 2P_{SOA} + P_{EDFA}/8)$
EPS + OSS	$N \times [n \times P_{TRx-élec} + (n+1) \times P_{EPS}]$	$N \times P_{Tx-accordable} + (P_{SOA} + 6 \times P_{O-E/E-O})/2$
POITR + PROTEUS	$N \times n \times (P_{TRx-accordable} + P_{BPF})$	$N \times (n \times P_{TRx-opt} + 4P_{OCS} + P_{WSS})$
EODCN	$N \times n \times (P_{QM} + P_{BVT} + 2 \times P_{TRx-opt}) + N \times (2 \times P_{VWSS} + P_{IAS} + P_{TRx-opt})$ $+N_p \times P_{OCS} + N \times (P_{TCAM} + P_{IES} + P_{TRx-opt}) + P_{SC}$	
POPI	$N \times n \times (P_{Tx-accordable} + P_{BPF}) + 2P_{ctl-ch} + 2P_{ctl-plane}$	

Tableau 2.5. Expressions de consommation d'énergie pour les différentes architectures d'interconnexion. N , n , k et c , étant le nombre de racks, de serveurs par rack, de commutateurs d'agrégation et cœur. La puissance consommée par les différents composants présents dans chacune de ces architectures est représenté par : $P_{TRx-élec}$ (transpondeurs électriques), P_{EPS} (EPS, consommation d'énergie par port), $P_{TRx-opt}$ (transpondeurs optiques à 10 Gbit/s), P_{OCS} (OCS, consommation d'énergie de la matrice de MEMS par port), P_{WSS} (WSS), P_{WC} (convertisseurs de longueur d'onde reconfigurables), P_{O-E}/P_{E-O} (conversions O-E/E-O), P_{SDRAM} (mémoires tampon), $P_{Tx-accordable}$ (émetteurs rapidement accordables en longueur d'onde), P_{SOA} (SOA), P_{EDFA} (EDFA), P_{BPF} (filtres passe-bande optiques), P_{QM} (files d'attente), P_{BVT} (transpondeurs de largeur de bande variable), P_{VWSS} (WSS de largeur de bande variable), P_{IAS} (structure d'ordonnement intra-rack), N_p (nombre de ports de la matrice à base de MEMS utilisé par EODCN), P_{TCAM} (TCAM), P_{IES} (structure d'ordonnement inter-racks), P_{SC} (structure de contrôle), P_{ctl-ch} (canal de contrôle), $P_{ctl-plane}$ (plan de contrôle). a correspond à la probabilité de contention pour DOS (0,2).

solution optique pour les connexions intra-baies, présentent des consommations d'énergie totales bien plus élevées que, par exemple, les architectures POITR + PROTEUS ou POPI.

2.4 Description du projet EPOC

Le projet EPOC est constitué de quatre sous-projets selon le niveau d'action sur le DC :

- **Réseau** : Design efficace d'une architecture interne d'interconnexion d'un centre de données à basse consommation d'énergie.
- **Infrastructure** : Design d'un système distribué en charge de l'optimisation de l'énergie

Consommation d'énergie (W)	
$P_{TRx-élec}$	0,3
$P_{TRx-opt}$	1
P_{EPS} par port	25
P_{OCS} par port	0,24
P_{WSS}	4
P_{WC}	1,5
P_{SDRAM}	0,9
$P_{Tx-accordable}$	3,5
P_{SOA}	2,5
P_{EDFA}	0,5
P_{BPF}	0,2
$P_{O-E/E-O}$	2,5
P_{QM}	3
P_{BVT}	8
P_{VWSS}	15
$P_{IAS/IES}$	2,5
P_{TCAM}	4,5
P_{SC}	300
P_{ctl-ch} par canal	1
$P_{ctl-plane}$	20

Tableau 2.6. Valeurs de consommation d'énergie des composants utilisés dans chaque architecture d'interconnexion.

consommée par l'infrastructure exécutant les tâches.

- **Application** : Conception d'une structure de gestion intelligente des ressources du *cloud* profitant de la disponibilité d'énergie renouvelable pour effectuer des tâches opportunistes et en explorant le compromis entre économie d'énergie et performance dans des systèmes distribués large échelle.
- **Modèles et programmation** : Contraintes de programmation et modèles de comportement formels.

Dans chaque sous-projet une thèse se déroule. Nos travaux ont été effectués dans le sous-projet « réseau ».

EPOC s'intéresse à la recherche sur les DC mono-site de petite et moyenne taille (entre 50 et 100 serveurs connectés). L'alimentation de ces serveurs est assurée via le réseau électrique courant et des ressources d'énergie renouvelable locales. EPOC propose la suppression du réseau de stockage (SAN), car il peut représenter jusqu'à 37% de l'énergie totale consommée par le DC. Des disques durs locaux sont alors mis en place au niveau de chaque serveur pour

remplacer le SAN.

Chaque serveur héberge des machines virtuelles (*VM : Virtual Machine*). Un hyperviseur gère les VM de tous les serveurs du DC. L'intérêt principal de la création des VM est de pouvoir regrouper les tâches qui se ressemblent en termes de ressources utilisées dans un même serveur physique. D'où, l'hyperviseur peut identifier un serveur qui exécute plusieurs tâches dont les ressources utilisées sont aussi utilisées par d'autres serveurs, le fait de migrer les VM de ce serveur vers d'autres serveurs nous permettrait d'éteindre ce premier et d'économiser de l'énergie. Ceci est connu comme politique VOVO (*Vary-On-Vary-Off*).

Donc, le design du DC doit prendre en compte la capacité de migrer toutes les VM d'un serveur (7,5 To soit la capacité totale d'un disque dur local). En considérant un port optique par rack, le débit sera partagé par les serveurs hébergés dans ce rack. Ainsi, est-ce que ce débit suffira pour la migration de toutes les VM d'un serveur en un temps suffisamment court ? En utilisant 10 Gbit/s par connexion, cette opération prendrait environ 1 heure et 40 minutes. Mais, si par exemple nous considérons 32 serveurs par rack, cette migration prendrait environ 53 heures et 20 minutes (parce que le débit de 10 Gbit/s serait partagé entre les 32 serveurs). En conséquence, il est nécessaire d'augmenter le débit de ces connexions avec une architecture de réseau optique qui ait une faible consommation d'énergie.

2.5 Conclusion

Dans ce chapitre nous avons présenté l'état de l'art des transmissions sur fibre optique pour les applications DC : les différents types d'émetteurs, récepteurs, formats de modulation, composants et technologies optiques. Nous avons vu les bases de la commutation optique et nous avons présenté des architectures d'interconnexion optiques pour DC. Nous avons identifié POPI, l'architecture d'interconnexion la plus efficace énergétiquement, que nous utiliserons pour réaliser les études des chapitres suivants. Nous avons enfin décrit le projet EPOC.

ÉTUDE DE L'ARCHITECTURE D'INTERCONNEXION POPI ET DE SON EXTENSION

3.1 Introduction

Ce chapitre est dédié à POPI, l'architecture d'interconnexion pour DC choisie pour le projet EPOC. Nous rappelons la performance de POPI en termes de latence moyenne et débit moyen. Ensuite, nous évaluons et comparons POPI en termes de bilan de liaison avec une autre architecture également basée sur TWIN, POITR. Puis, nous proposons des variantes de l'architecture POPI originale avec d'autres options de connexion des canaux de contrôle au coupleur central. Enfin, nous introduisons des extensions regroupées sous le terme d'E-POPI, utilisant plusieurs bandes optiques et des lasers accordables chacun dans une des bandes.

3.2 Analyse de POPI

3.2.1 Evaluation de la performance

Jusqu'à présent, aucune des références traitant sur les architectures d'interconnexion tout-optiques pour DC (présentées dans le chapitre précédent), à l'exception de POPI, n'a considéré pour l'étude de sa performance un modèle de trafic adapté aux DC. Le simulateur de POPI prend en compte l'étude présentée par Benson et al. [96]-[97]. Cette étude porte sur différents types de DC, en termes de nombre de dispositifs utilisés dans les couches cœur, d'agrégation, ToR et serveurs. Le trafic d'un DC correspond à environ 80% à des transferts à l'intérieur d'un rack (dont la plupart du trafic de fond) et 20% aux échanges entre les différents racks du DC (principalement du trafic lourd).

Les auteurs de [86] ont implémenté leur simulateur de réseau sous le logiciel *Omnet++* [98], en prenant en compte deux types de trafic :

- des « souris », qui correspondent à des rafales de données qui doivent être traitées en temps réel

- des « éléphants », qui correspondent à des grandes quantités des données à transmettre pas forcément à temps réel

Ce simulateur permet de calculer la latence moyenne pour le trafic de type « souris » et le débit moyen pour le trafic de type « éléphant » en fonction de la charge ρ_w . Il considère deux types de flux : « *greedy* », qui correspond au trafic soit de type « souris » soit de type « éléphant » ; et « *query* », à l'ensemble de paquets de demande et d'autorisation de transmission (*report* et *grant*). Le premier est représenté par des arrivées de paires source-destination selon le processus de Poisson. Ce flux correspond à un nombre donné de paquets de taille 1 ko distribué géométriquement avec une taille moyenne de 10 Mo. Le trafic est généré symétriquement par 16 ou 64 sources indépendantes à un débit maximal C de 1 Gbit/s. Une bande de garde Δ_g de 200 μ s entre chaque rafale est prise en compte. Le deuxième est caractérisé par des paquets de 2 o et de 10 o pour les messages de *report* et de *grant*, respectivement. Le débit du *report* et du *grant* est également de 1 Gbit/s.

La variable Q , introduite par les auteurs de [86], est utilisée comme unité d'allocation (en octets). Les flux de type *greedy* peuvent avoir en attente des paquets de taille supérieure à Q ; mais, les flux de type *query* peuvent, quant à eux, n'avoir en attente que des paquets de taille inférieure ou égale à Q . Cette variable est utilisée pour exprimer la latence moyenne et le débit moyen en fonction de la charge ρ_w , selon les expressions suivantes :

$$\text{Latence moyenne} \approx 257 \mu\text{s} + \frac{Q + C \times \Delta_g}{2 \times C \times \rho_w} \times (1 - \rho_w) \quad (3.2.1)$$

$$\text{Débit moyen} = \frac{C \times Q}{Q + C \times \Delta_g} \times (1 - \rho_w) \quad (3.2.2)$$

Où 257 μ s correspond au temps de signalisation qui est égale au temps d'attente moyen entre chaque flux de type *query* (250 μ s) plus les temps de propagation (environ 1 μ s par tranche de 200 m de fibre optique) du *report* et du *grant*. Le terme $\rho_w/(1 - \rho_w)$ représente le nombre moyen de flux en attente qui est égal au nombre de clients qui partagent la queue.

La Figure 3.1 montre les résultats de (a) latence moyenne et de (b) débit moyen en fonction de la charge présentés dans [86].

La valeur minimale de la latence moyenne est déterminée par le temps de signalisation. La latence reste proche de cette valeur moyenne lorsque la charge est suffisamment faible devant 1. La latence est d'autant plus proche de la valeur minimale que le nombre de sources augmente. Le débit moyen décroît naturellement avec la charge, de façon inversement proportionnelle dans la limite d'un nombre infini de sources.

Nous pouvons nous appuyer sur les résultats de cette étude pour estimer les performances de POPI dans le contexte du projet EPOC. En effet, les hypothèses de trafic présentées en [86] sont proches de celles d'EPOC : un trafic de fond (de type « souris ») et un trafic lourd contenant toutes les VM d'un serveur (de type « éléphant »). Bien que le débit pris en compte

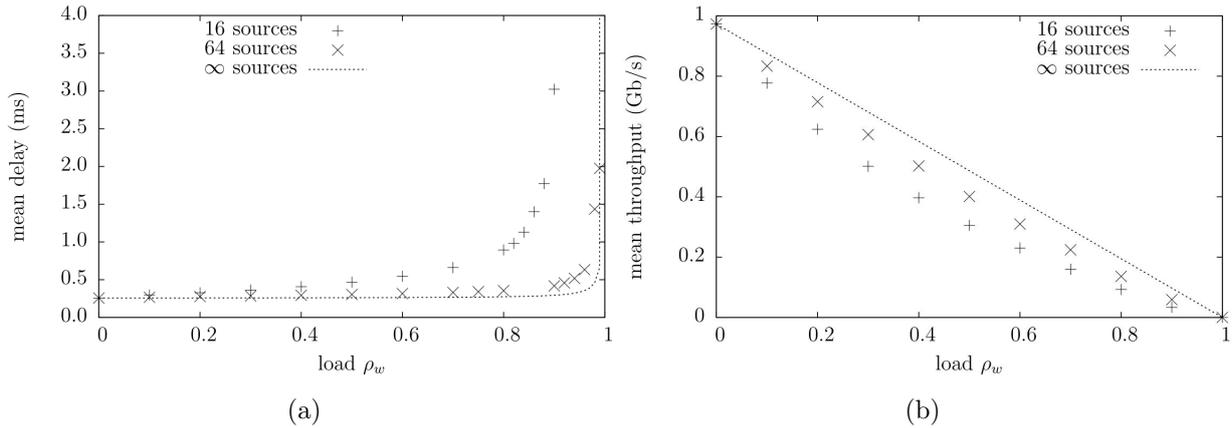


Figure 3.1. Performance de POPI en termes de : (a) latence moyenne ; (b) débit moyen par serveur en fonction de la charge ρ_w [86].

dans les simulations précédentes soit d'1 Gbit/s au lieu de 100 Gbit/s dans le projet EPOC, ces résultats peuvent être considérés comme toujours pertinents à un facteur 100 près dans l'échelle de temps. Toutefois, la contribution du temps de propagation à la latence minimale n'est plus négligeable (de l'ordre d'1 μ s comparé à 2 μ s entre deux messages *query*).

3.2.2 Bilan de liaison

Comme déjà mentionné dans les paragraphes précédents, nous cherchons à atteindre un débit de 100 Gbit/s. Le fait d'augmenter le débit va rendre le signal plus sensible à la dispersion chromatique. Une façon (la plus simple) d'améliorer le taux d'erreurs binaire (*BER : Bit Error Rate*) de ce signal est d'augmenter sa puissance à l'émission.

Toutefois, il ne suffit pas de seulement augmenter le débit. Sous le format modulation de base NRZ, par exemple, le signal requis aurait une bande spectrale d'environ 150 GHz si nous considérons 75% du premier lobe (l'équivalent de trois canaux ITU espacés de 50 GHz). Ainsi, ceci pourrait être obtenu par l'utilisation de formats de modulation multiniveaux spectralement plus efficaces (selon le nombre de niveaux de modulation), cependant requièrent une puissance reçue plus élevée afin de compenser l'interférence intersymboles.

En conséquence, cette section est consacrée à l'étude du bilan de liaison de POPI¹ (selon le schéma de la Figure 2.26), à estimer les pertes de chaque composant de cette architecture pour pouvoir déterminer une certaine marge pour la puissance optique nécessaire en réception. Par ailleurs, nous étudions de même le bilan de liaison de POITR (selon les schémas (a) et (b) de la Figure 2.23). Nous voulons comparer POPI avec POITR et décider laquelle de ces deux architectures d'interconnexion serait la plus adaptée au projet EPOC de ce point de vue.

Les bilans de liaison de POPI, POITR (a) et (b) sont décrits selon les équations 3.2.3,

1. Nous prenons en compte l'option 2 de POPI (1 longueur d'onde par serveur).

3.2.4 et 3.2.5, respectivement.

$$\check{P}_{Tx} \geq P_{Rx} + L_{fibre} + L_{connecteurs} + L_{EAM} + L_{BPF} + L_{C(total)} + M \quad (3.2.3)$$

$$\begin{aligned} \check{P}_{Tx} \geq & P_{Rx} + L_{fibre} + L_{connecteurs} + L_{EAM} + L_{BPF} + L_{C(total),A(a)} \\ & + L_{circulateur} + L_{FBG} + L_{C(total),B90\%(a)} + L_{AWG1} + M \end{aligned} \quad (3.2.4)$$

$$\begin{aligned} \check{P}_{Tx} \geq & P_{Rx} + L_{fibre} + L_{connecteurs} + L_{EAM} + L_{BPF} + L_{circulateur} \\ & + L_{FBG} + L_{C(total),A(b)} + L_{AWG1} + M \end{aligned} \quad (3.2.5)$$

Où \check{P}_{Tx} est la puissance de transmission optique minimale qui assure un BER $\leq 10^{-12}$ à 10 Gbit/s sous le format de modulation NRZ sans FEC (dans un premier temps nous ne considérons pas de FEC pour des raisons de coût et de latence ; dans un deuxième temps, en augmentant le débit, il devient nécessaire en dépit du coût et de la latence). P_{Rx} représente la sensibilité du récepteur ; L_{fibre} , les pertes de la fibre optique ; $L_{connecteurs}$, les pertes des connecteurs ; L_{EAM} , les pertes du modulateur (électro-absorbant) ; L_{BPF} , les pertes du filtre passe-bande en réception ; $L_{C(total)}$, les pertes du coupleur étoile de POPI (selon le schéma de la Figure 2.26) et M , une marge de 1,5 dB. L_{FBG} , $L_{C(total),A(a)}$ et $L_{C(total),B90\%(a)}$ correspondent aux pertes liées au filtre de Bragg, aux coupleurs A et B (90%) de POITR (a) (selon le schéma de la Figure 2.23(a)), respectivement. L_{AWG1} , $L_{circulateur}$ et $L_{C(total),A(b)}$ sont les pertes dues au AWG1, au circulateur et au coupleur de POITR (b) (selon le schéma de la Figure 2.23(b)), respectivement.

Nous faisons cette comparaison pour deux types de récepteurs différents : PIN et APD, pour chaque architecture. Nous prenons en compte des valeurs de sensibilité de -19 et -24 dBm [83], respectivement, pour un BER à 10^{-12} (à ce stade, nous ne considérons pas de FEC).

Pour estimer les pertes d'insertion du coupleur passif $x : x$ (avec $x = 2n + 3$) de POPI, $L_{C(total)}$, et celles du coupleur passif $1 : x$ (avec $x = n$) de POITR, $L_{C(total),A(a)}$ ou (b) , en fonction de x , nous considérons :

$$L_{C(total)} = L_{C(total),A(a) \text{ ou } (b)} = L_C + L_{C,excès} \quad (3.2.6)$$

Où L_C correspond aux pertes du coupleur dus au nombre de ports connectés x :

$$L_C = 10 \times \log(x) \quad (3.2.7)$$

$L_{C,excès}$ représente les pertes d'excès du coupleur, que nous estimons grâce à l'équation suivante :

$$L_{C,excès} = (L_C + 3)/6 \quad (3.2.8)$$

Nous avons établi cette expression à partir de plusieurs spécifications de produits aujourd'hui disponibles dans le commerce [99]-[101].

Pour les autres composants nous utilisons les pertes indiquées dans le Tableau 3.1.

Composant	Pertes (dB)	Référence
L_{fibre} (1 km)	0,2	[102]-[103]
$L_{connecteurs}$	1	[104]
L_{EAM}	3,3	[105]
L_{BPF}	5	[106]
$L_{circulateur}$	0,6	[83],[107]
L_{FBG}	0,4	[83],[108]
$L_{C(total),B90(a)}$	0,65	[83]
L_{AWG1}	2,5	[83]

Tableau 3.1. Pertes des composants à considérer pour le bilan de liaison des architectures POPI et POITR options (a) et (b).

La Figure 3.2 montre la puissance d'émission minimale par canal \check{P}_{Tx} en fonction du nombre de serveurs connectés (pour le réseau POPI trois longueurs d'onde ont été réservées pour module de contrôle et le *gateway*).

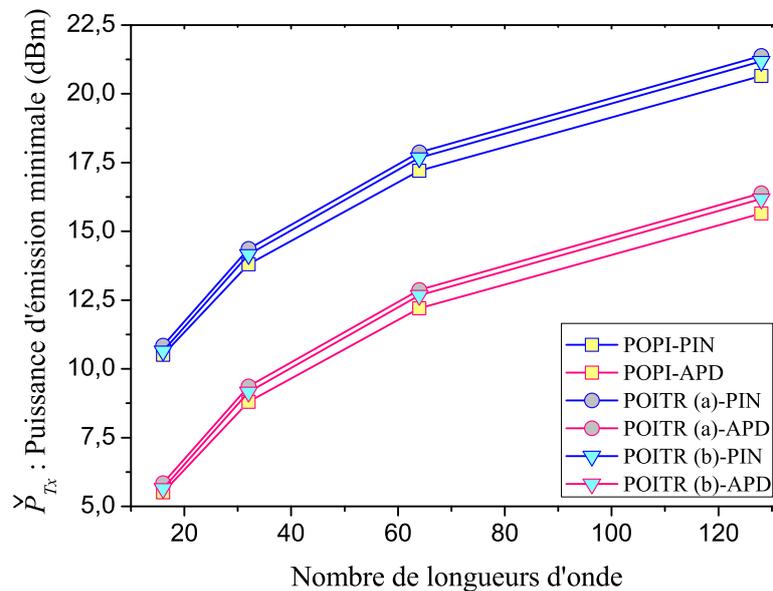


Figure 3.2. Comparaison du bilan de liaison entre POPI et POITR.

D'après ce graphique nous pouvons constater que la puissance d'émission minimale \check{P}_{Tx} de l'architecture POPI est toujours au-dessous de celle des deux options de POITR. Ceci est un résultat attendu, puisque POITR (a) et (b) utilisent des composants passifs en plus, qui augmentent ainsi les pertes totales du bilan de liaison. A la base, les auteurs de POITR

proposent l'option (b) afin de réduire les pertes de puissance totales, cependant, grâce à la Figure 3.2, nous constatons que cette baisse est négligeable par rapport aux résultats de POPI. D'autre part, nous observons que, conformément au chapitre précédent, \check{P}_{Tx} est toujours plus faible pour les récepteurs APD.

Nous en déduisons que POPI est le meilleur choix, en termes de puissance d'émission minimale. D'autre part, nous considérons que la puissance d'émission minimale de POPI pourrait être diminuée si nous ne connectons pas les transpondeurs de contrôle directement au coupleur passif. En conséquence, nous proposons ci-dessous des variantes de POPI qui permettent de faire des économies en ce qui concerne le bilan de liaison.

3.2.3 Variantes de POPI

A partir de maintenant, nous allons nous référer à la structure de POPI que nous avons présentée dans le Chapitre 2 (Figure 2.26) comme « POPI option A ». Dans cette section, sur les figures 3.3, 3.4 et 3.5 nous présentons trois nouvelles options (B, C et D). Pour ces trois options, le fait de ne pas connecter les transpondeurs de contrôle directement au coupleur passif $x : x$ nous permet de considérer $x = n + 3$ (au lieu de $x = 2n + 3$). Autrement dit, nous allons pratiquement doubler le nombre de serveurs connectés au coupleur.

L'option B considère les longueurs d'onde de données dans les bandes de transmission C ou L et celles de contrôle (λ_r et λ_g) dans la bande de transmission O, ayant pour but de pouvoir les séparer en émission et en réception à l'aide d'un duplexeur à faibles pertes d'insertion (0,5 dB) [109]. L'interface et la structure du réseau de l'option B sont présentés sur les figures 3.3(a) et (b), respectivement. Ainsi, nous devons ajouter 1 dB de pertes liées au duplexeur (en émission et en réception) dans le bilan de liaison de l'équation 3.2.3.

L'option C considère les longueurs d'onde de données et de celle du *report* (λ_r) dans les bandes de transmission C ou L et celle de *grant* (λ_g) dans la bande de transmission O. Cette fois-ci l'émetteur accordable de POPI va pouvoir configurer les longueurs d'onde des données et du *report*. Le duplexeur sera utilisé seulement en réception pour séparer les longueurs d'onde des données de celle du *grant*. L'interface et la structure du réseau de l'option C sont présentés sur les figures 3.4(a) et (b), respectivement. Nous devons ajouter 0,5 dB de pertes du duplexeur (en réception) dans le bilan de liaison de l'équation 3.2.3.

L'option D considère les longueurs d'onde de données et celles de contrôle (λ_r et λ_g) dans la même bande de transmission (C ou L). Encore une fois l'émetteur accordable de POPI va pouvoir configurer les longueurs d'onde des données et du *report*. Les longueurs d'onde des données vont être séparées de celle du *grant* en réception à l'aide d'un filtre de Bragg et d'un circulateur. L'interface et la structure du réseau de l'option D sont présentés sur les figures 3.5(a) et (b), respectivement. Nous devons ajouter 1 dB de pertes (0,6 et 0,4 dB liées au circulateur et au filtre de Bragg, respectivement, Cf. Tableau 3.1) dans le bilan de liaison de l'équation 3.2.3.

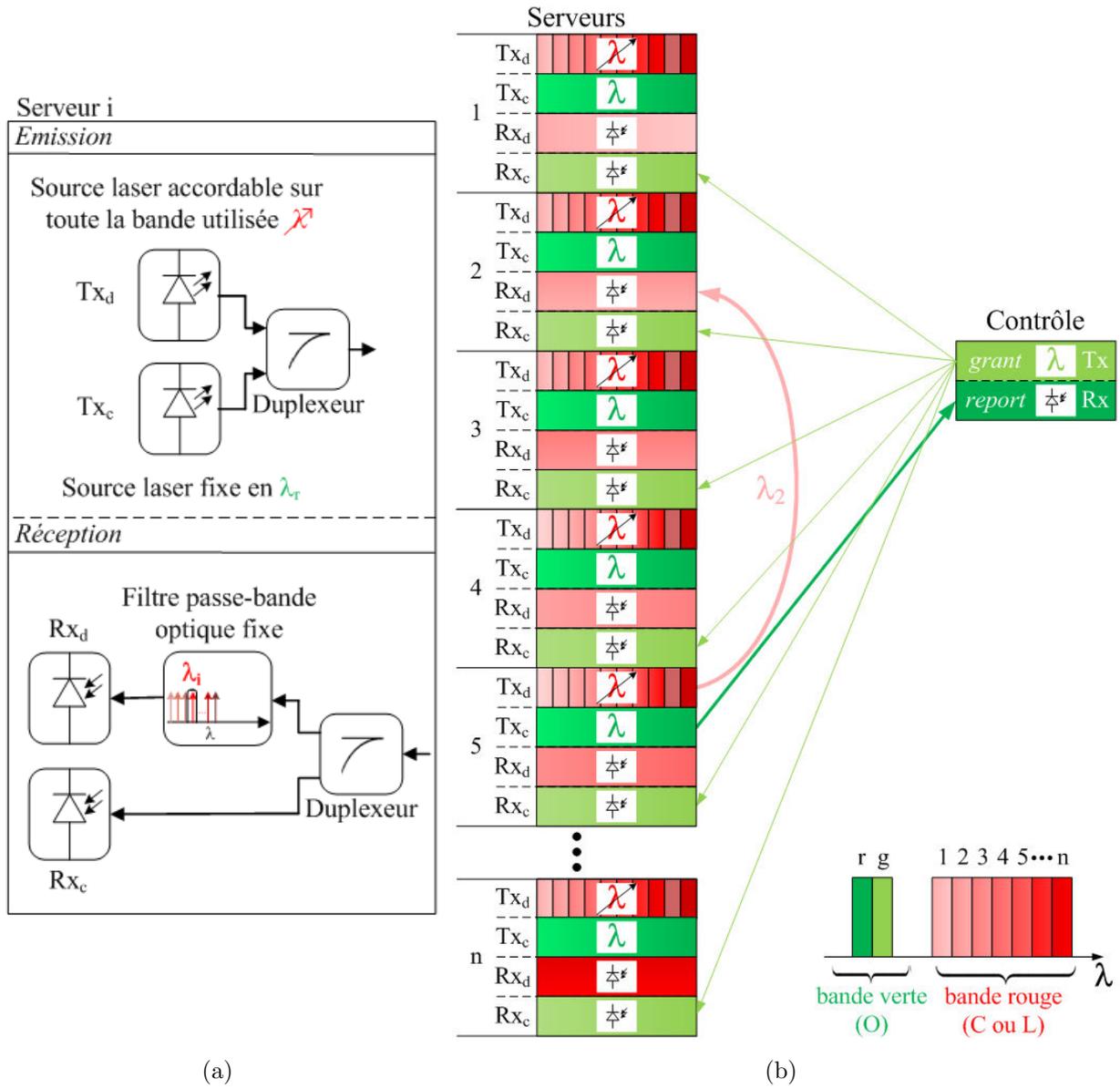


Figure 3.3. POPI option B : (a) interfaces de données et de contrôle. (b) exemple de connexion entre les serveurs 5 et 2, Tx_d et Tx_c représentent les émetteurs de données et de contrôle ; Rx_d et Rx_c , les récepteurs de données et de contrôle, respectivement.

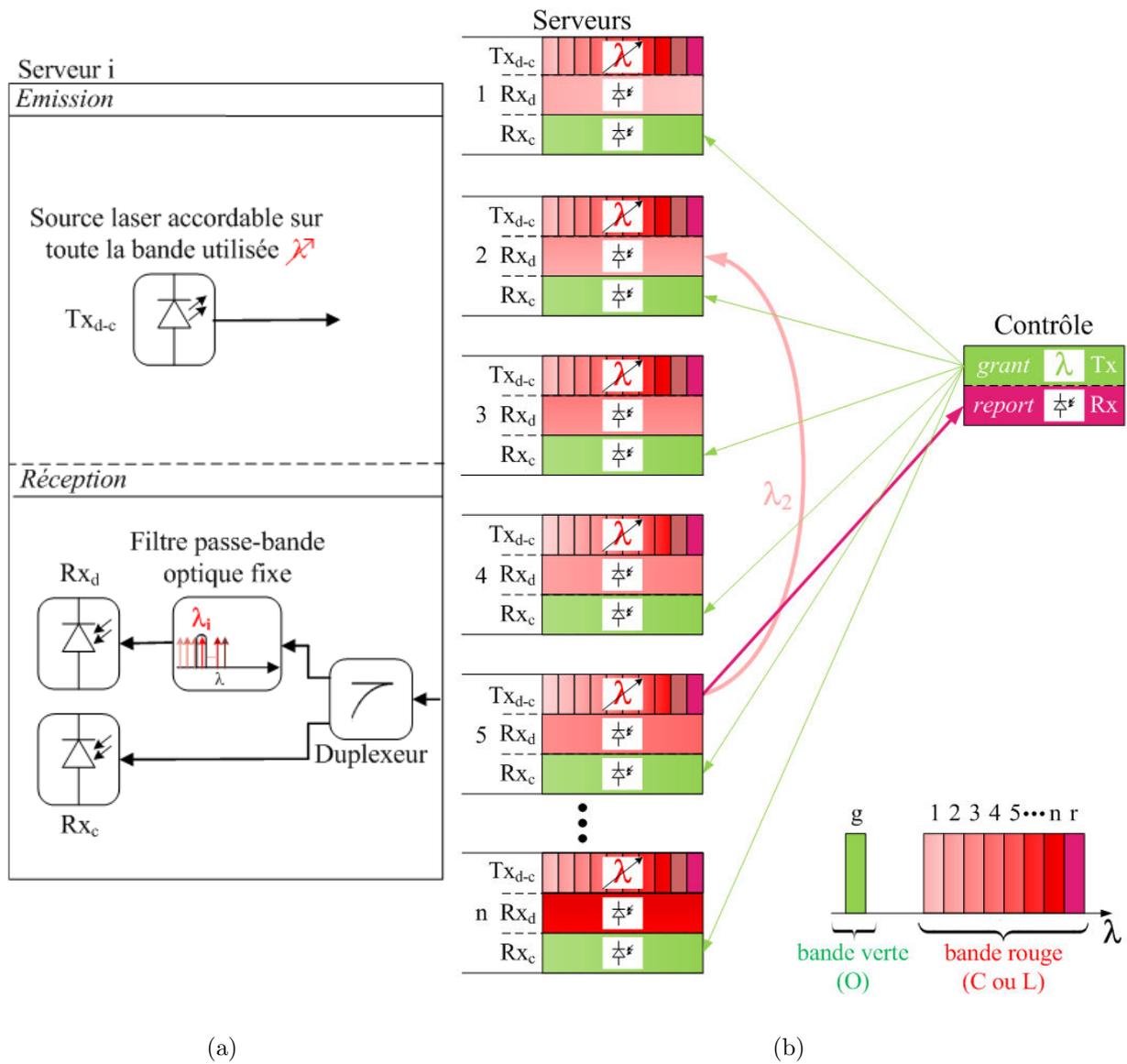


Figure 3.4. POPI option C : (a) interfaces de données et de contrôle. (b) exemple de connexion entre les serveurs 5 et 2, Tx_d et Tx_c représentent les émetteurs de données et de contrôle ; Rx_d et Rx_c , les récepteurs de données et de contrôle, respectivement.

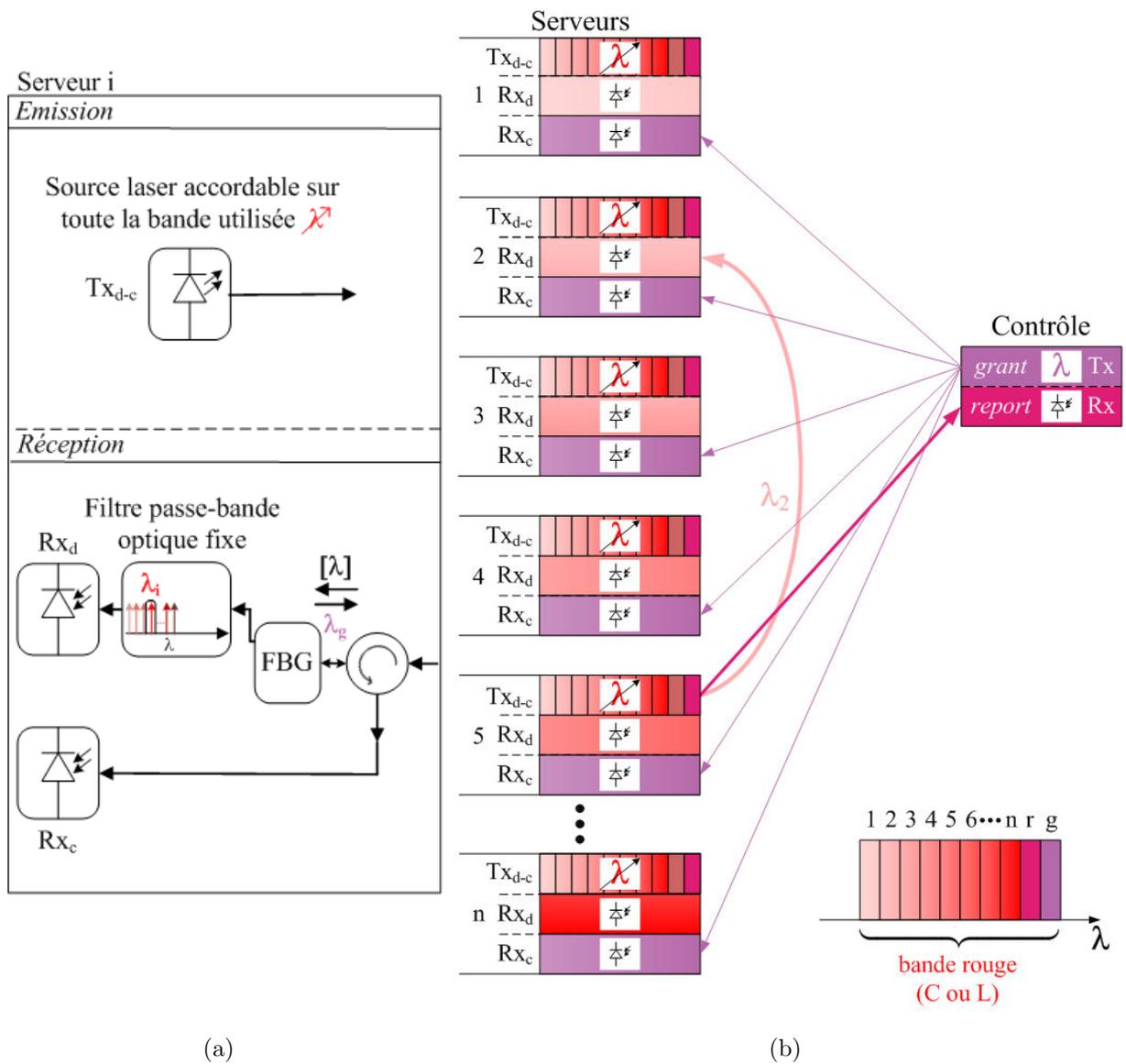


Figure 3.5. POPI option D : (a) interfaces de données et de contrôle. (b) exemple de connexion entre les serveurs 5 et 2, Tx_d et Tx_c représentent les émetteurs de données et de contrôle ; Rx_d et Rx_c , les récepteurs de données et de contrôle, respectivement.

La Figure 3.6 montre la puissance d'émission minimale \check{P}_{Tx} en fonction du nombre de serveurs connectés pour les quatre options de POPI. Effectivement, les options B, C et D requièrent des puissances d'émission minimales bien moins élevées que l'option originale (A). Cette valeur pourrait être encore améliorée en utilisant des récepteurs optiques ayant une sensibilité supérieure (par exemple avec une pré-amplification optique, cas étudié en détail dans le Chapitre 5).

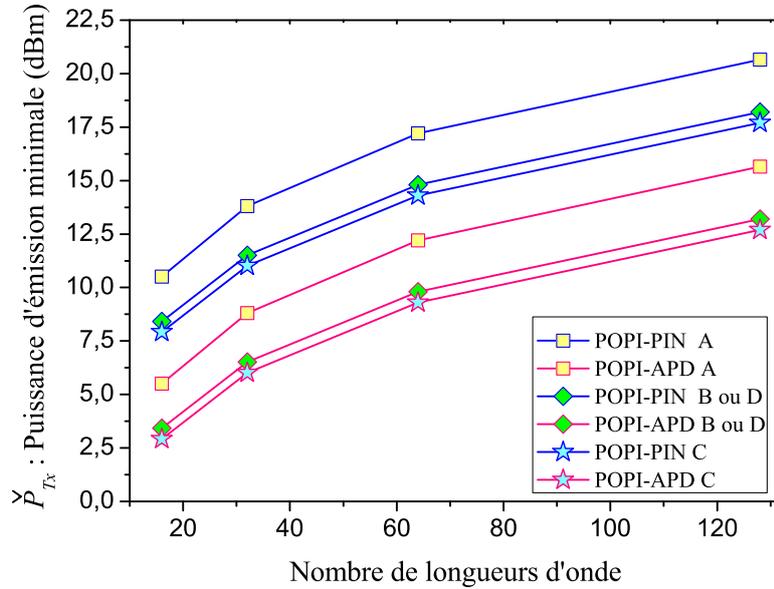


Figure 3.6. Comparaison du bilan de liaison entre les options A, B, C et D de POPI.

3.2.4 Limite de capacité

La capacité maximale de POPI en termes de nombre de serveurs connectés (avec une longueur d'onde dédiée par serveur) est liée aux contraintes imposées par chaque composant de l'architecture. D'une part, elle est limitée par les caractéristiques de l'émetteur rapidement accordable en longueur d'onde et de son électronique. La plage d'accord et l'espacement entre canaux vont déterminer le nombre maximal de serveurs connectés permis par l'émetteur. En outre, la vitesse d'accord des lasers augmente au détriment de la précision de la longueur d'onde accordée, compromis à considérer pour le choix de l'émetteur. D'autre part, la capacité de POPI est limitée par la largeur de bande de chaque canal multiplexé en longueur d'onde et par la disponibilité des filtres optiques passe-bande correspondants. En principe, la largeur de bande du canal peut être réduite de manière significative en augmentant la complexité du format de modulation.

Accroître la complexité du format de modulation est un levier essentiel pour améliorer les performances de POPI. Soit, à largeur de canal constante, pour augmenter le débit de chaque canal ; soit, à débit constant pour augmenter le nombre de canaux et donc de serveurs connectés (Figure 3.7).

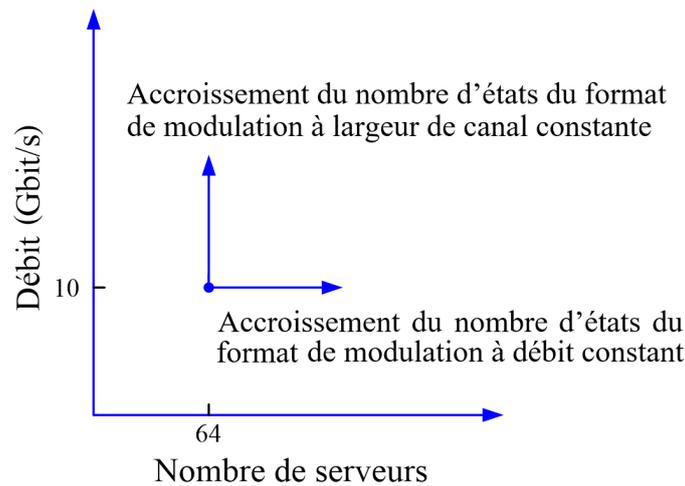


Figure 3.7. Débit de chaque canal en fonction du nombre de serveurs dans le DC, dans le cas où la bande totale est limitée par la source accordable.

Notons que des formats encore plus complexes que ceux qui sont actuellement à l'étude pour les applications aux télécommunications à longue distance pourraient être envisagés, dû au fait que les défauts de propagation linéaires et non-linéaires sont considérablement réduits pour des courtes distances. Ce qui nous permet d'augmenter encore la puissance d'émission du canal, afin d'atteindre un OSNR supérieur au minimum requis par ces formats de modulation. Des questions pratiques, telles que : le coût, la sécurité oculaire ou la consommation d'énergie, doivent être pris en compte lors de la définition d'un compromis approprié.

Dans notre étude nous prendrons comme référence l'émetteur accordable en longueur d'onde présenté par Simsarian et al. [26]-[27], avec 64 canaux espacés de 50 GHz dans une bande d'accord de 3200 GHz. Son temps d'accord maximal est de 50 ns. Cet émetteur nous permettrait ainsi de connecter 63 serveurs à POPI options A et B, tout en réservant une longueur d'onde de *gateway*. L'option C de son côté peut connecter jusqu'à 62 serveurs, car elle doit encore réserver une longueur d'onde de *report* dans la bande du laser accordable. Finalement, l'option D permet la connexion de 61 serveurs, parce qu'il faudrait encore réserver une longueur d'onde de *grant*.

3.3 POPI étendu : E-POPI

Afin d'augmenter le nombre d'entités connectées à POPI, limité par la bande de transmission du laser accordable en longueur d'onde, dans cette section nous proposons une extension de POPI que nous appelons E-POPI. Ce nouveau réseau considère plusieurs émetteurs accordables en longueur d'onde dans b bandes de transmission différentes. De cette façon, cette extension permet en principe d'augmenter d'un facteur b , le nombre de serveurs connectés à POPI.

Nous avons déjà vu que les options B, C ou D de POPI permettent de limiter les pertes

d'insertion du coupleur. Bien que l'option B permette de mieux utiliser la bande de transmission du laser accordable et que l'option C présente les pertes d'insertion les plus faibles, ces deux options nécessitent d'utiliser une bande spécifique pour le canal de contrôle et un duplexeur. Ce dernier existe aujourd'hui dans le commerce pour des canaux de contrôle en bande O et pour des canaux de données en bande C ou L, mais pas C+L comme nous aurons besoin pour E-POPI. De son côté, l'option D utilise des composants disponibles dans le commerce depuis longtemps (circulateur [107] et FBG [108]). De plus, dans cette option, les deux canaux de contrôle sont dans la même bande de transmission que les canaux de données, ce qui est fortement souhaitable pour utiliser des bandes supplémentaires. Pour ces raisons, les structures d'E-POPI que nous présentons par la suite sont basées sur POPI option D.

Nous parlons ci-dessous de structure $C - T - F - R$. C représente le nombre de transpondeurs (Tx/Rx) au niveau du contrôleur. T est le nombre d'émetteurs de données (Tx_d) par serveur (nombre d'émetteurs accordables en longueur d'onde). F correspond au nombre de filtres optiques passe-bande par récepteur de données par serveur ; R , au nombre de récepteurs de données (Rx_d) par serveur. Ainsi, la capacité potentielle d'E-POPI augmente d'un facteur $b = T \times F \times R$ par rapport au réseau POPI option D.

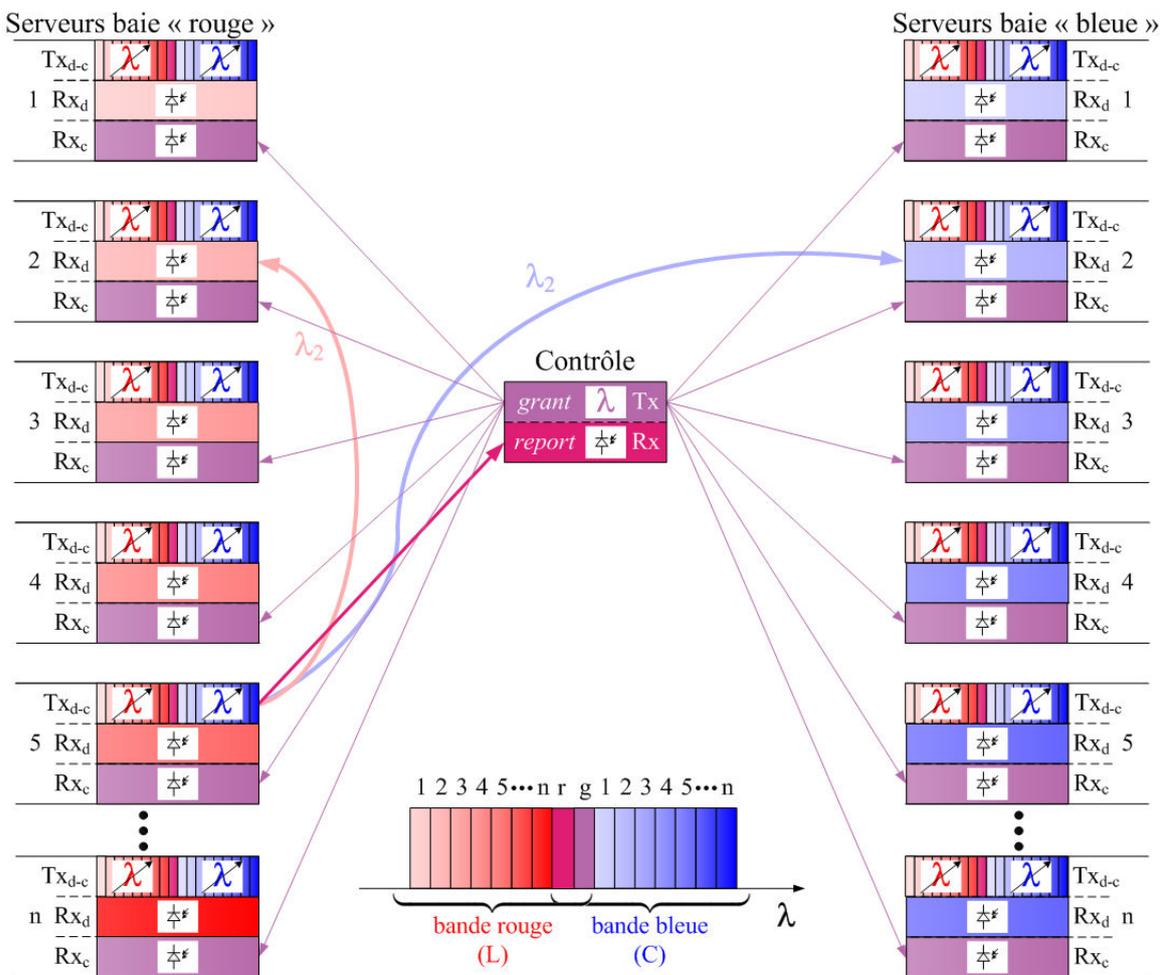
3.3.1 Options d'E-POPI à deux bandes

Ces options considèrent deux baies ou ensembles de serveurs connectés (rouge et bleu). Le principe consiste à positionner les longueurs d'onde de *report* et de *grant* entre les deux bandes de transmission optiques que nous désirons accoler. Cela suppose que les plages des émetteurs accordables en longueur d'onde se chevauchent sur ces canaux. Nous allons ensuite doubler soit le nombre d'émetteurs accordables, soit celui de filtres optiques ou celui de photodiodes en réception. Le choix du nombre d'émetteurs et de récepteurs par serveurs aura naturellement un impact sur la charge admissible par serveur.

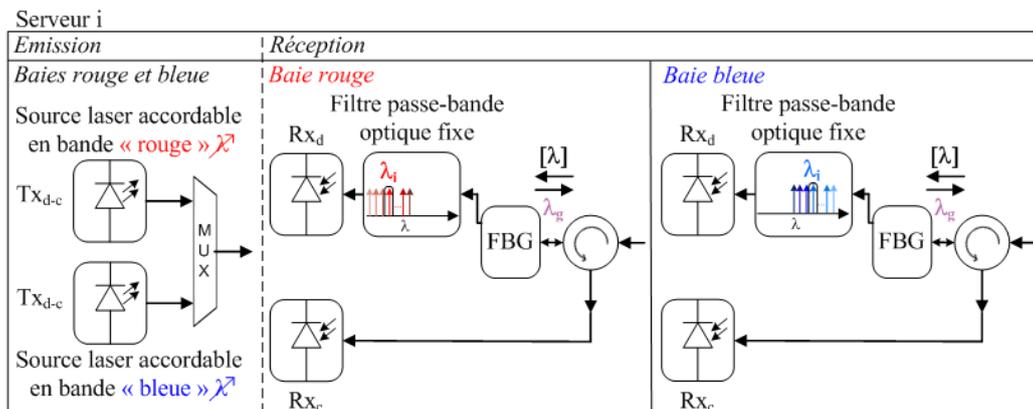
3.3.1.1 E-POPI 1-2-1-1

Chaque serveur est équipé de deux émetteurs accordables en longueur d'onde sur des bandes de transmissions optiques adjacentes (le choix de l'émetteur correspond à la baie de destination) et d'un récepteur unique (un filtre passe-bande optique en réception différent pour chaque serveur et chaque baie, une seule photodiode). La Figure 3.8(a) montre le schéma d'interconnexion d'E-POPI 1 – 2 – 1 – 1 ; la Figure 3.8(b), les structures de l'émetteur et du récepteur.

La structure du contrôleur ne change pas, car nous n'utilisons qu'une seule paire de Tx/Rx pour l'envoi des messages de *report* et de *grant*. La complexité du plan de contrôle n'est accrue que légèrement par rapport à POPI, puisque il faut maintenant préciser lequel des émetteurs il faut accorder selon la baie et le serveur de destination. La charge maximale autorisée n'est pas modifiée par rapport à POPI car chaque serveur a un filtrage et une



(a)



(b)

Figure 3.8. E-POPI 1 – 2 – 1 – 1 : (a) schéma d’interconnexion d’E-POPI 1 – 2 – 1 – 1, 1, 2, ..., n : longueurs d’onde relatives aux serveurs 1, 2,..., n, respectivement, r et g : longueurs d’onde relatives aux messages de *report* et de *grant*, respectivement ; (b) structure en émission et en réception.

photodiode uniques. La possibilité de diffusion n'est pas envisageable. Le bilan de liaison et la consommation d'énergie varient par rapport à ceux de POPI option D. L'utilisation d'un multiplexeur de bande en émission augmente les pertes du réseau de 1 dB [110]-[111]. La consommation d'énergie de cette option augmente donc de 3,5 W par serveur, correspondant à l'émetteur accordable en longueur d'onde additionnel.

3.3.1.2 E-POPI 1-1-2-1

Chaque serveur est équipé d'un seul émetteur accordable en longueur d'onde avec une bande de transmission propre à sa baie (par exemple, bande C pour la baie rouge et bande L pour la baie bleue) et de deux filtres optiques passe-bande pour un seul récepteur de données. Ces filtres sont centrés à la longueur d'onde relative à l'indicateur du serveur (pour les deux bandes de transmission optiques). Le schéma d'interconnexion de cette option est présenté sur la Figure 3.9(a) ; les structures de l'émetteur et du récepteur, sur la Figure 3.9(b).

La structure du contrôleur ne change pas, puisque nous utilisons toujours une seule paire de Tx/Rx pour l'envoi des messages de *report* et de *grant*. La complexité du plan de contrôle est accrue, car cette option fonctionne avec un seul récepteur de données et ne peut traiter qu'un seul flux provenant des bandes de transmission C ou L à la fois. Ainsi, E-POPI 1 – 1 – 2 – 1 suppose de pouvoir mettre en mémoire tampon l'un des deux flux (flux en bande C ou L destiné au même serveur en même temps) afin d'éviter les collisions.

La Figure 3.9(a) montre un exemple de connexion entre les serveurs 5 et 2 appartenant à la baie rouge. Le fait d'avoir équipé chaque serveur avec deux filtres en réception (pour pouvoir recevoir les paquets provenant des deux différentes baies) sous une architecture de diffusion comme POPI, permet au serveur 2 de la baie bleue de recevoir ce paquet également. Ceci pose des problèmes de sécurité des données transférées ou de possibles collisions si jamais un autre serveur essaye d'envoyer un autre paquet au serveur 2 de la baie bleue (paquet qui serait également diffusé vers le serveur 2 de la baie rouge et qui entrerait en collision avec celui envoyé par le serveur 5 de la baie rouge). Le conflit de la sécurité des données doit être résolu en ajoutant un certain cryptage (par exemple, le cryptage G983/G984 utilisé dans les PON [112]). Les collisions peuvent être évitées en adaptant le plan de contrôle de POPI.

Il faut tenir compte du fait que cette option va doubler le nombre de serveurs sans doubler sa capacité de traitement de flux/paquets, car le fait de n'avoir qu'un seul récepteur de données par serveur empêchera celui-ci de traiter deux flux en parallèle. Ceci se répercute sur la charge maximale admissible, qui diminue de 50% par rapport au réseau POPI. La possibilité de diffusion d'E-POPI 1 – 1 – 2 – 1 pourrait être éventuellement exploitée en configurant le DC de telle sorte que les couples de serveurs les plus susceptibles de devoir s'en servir soient logés dans des baies différentes et sous le même indicateur de serveur (positionnement de la longueur d'onde en réception dans la bande de transmission respective). Le bilan de liaison et la consommation d'énergie varient par rapport à ceux de POPI option D. L'utilisation de deux (dé)multiplexeurs de bande en réception augmente les pertes du

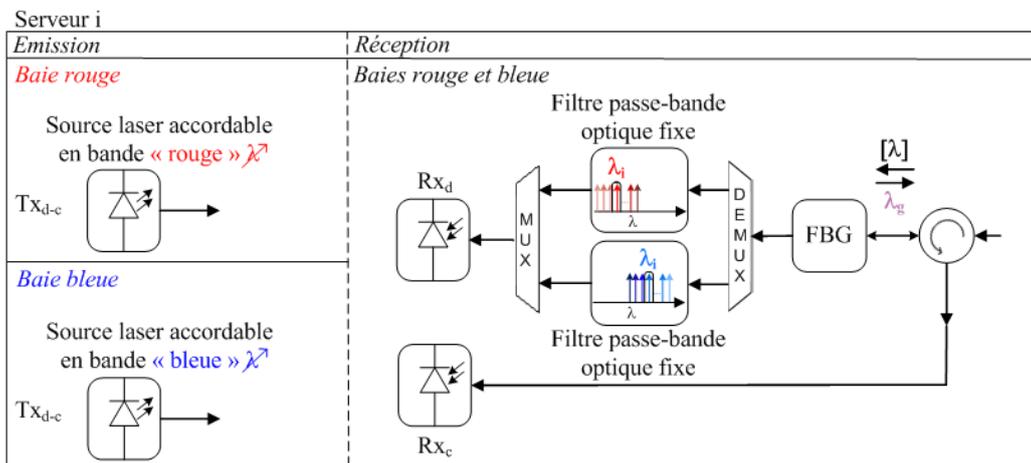
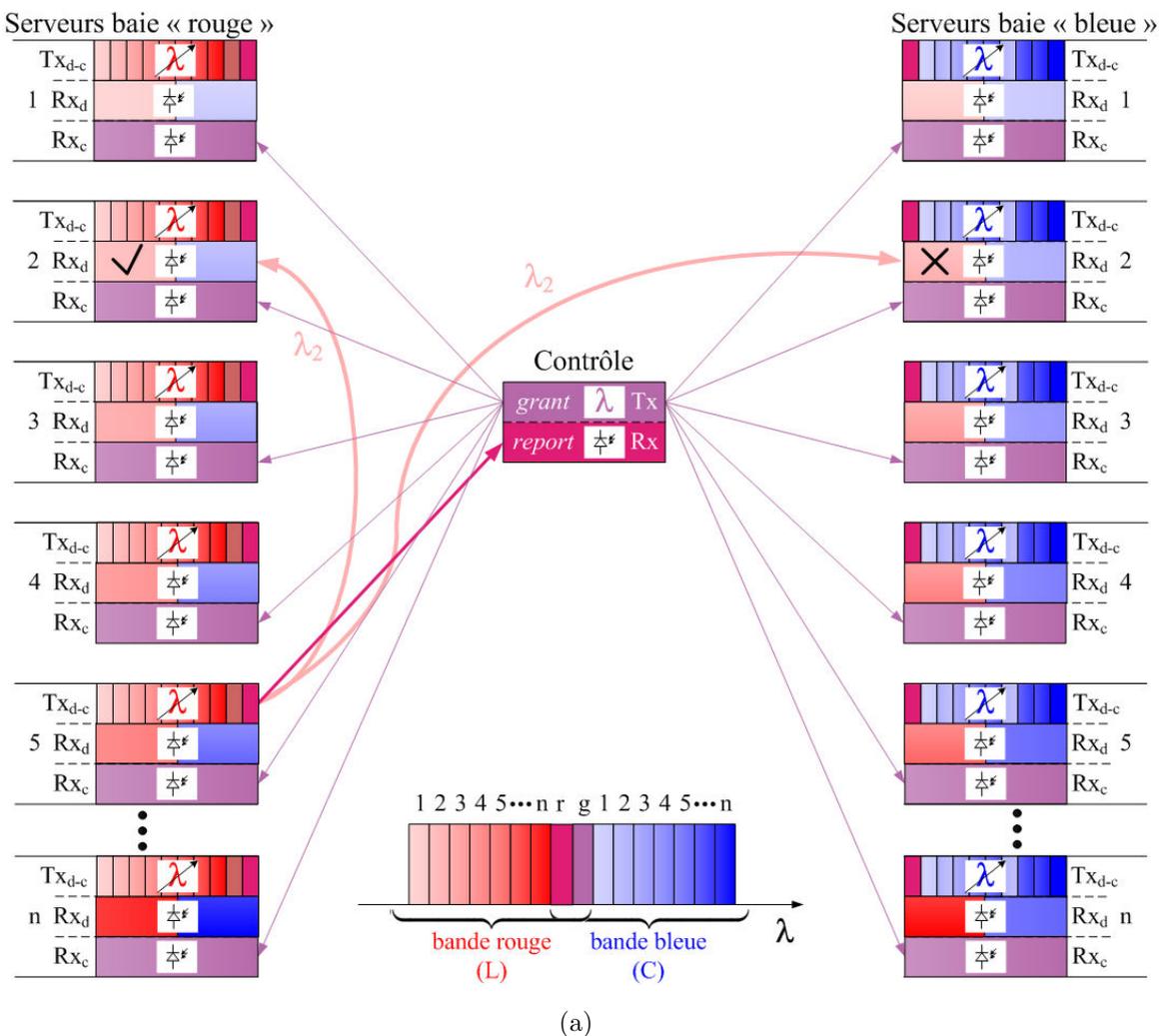


Figure 3.9. E-POPI 1 – 1 – 2 – 1 : (a) schéma d’interconnexion, 1, 2, ..., n : longueurs d’onde relatives aux serveurs 1, 2,..., n, respectivement, r et g : longueurs d’onde relatives aux messages de *report* et de *grant*, respectivement ; (b) structures en émission et en réception.

réseau de 2 dB. La consommation d'énergie de cette option augmente de 0,2 W par serveur, correspondant au filtre optique passe-bande (bande C ou L) additionnel.

3.3.1.3 E-POPI 1-1-1-2

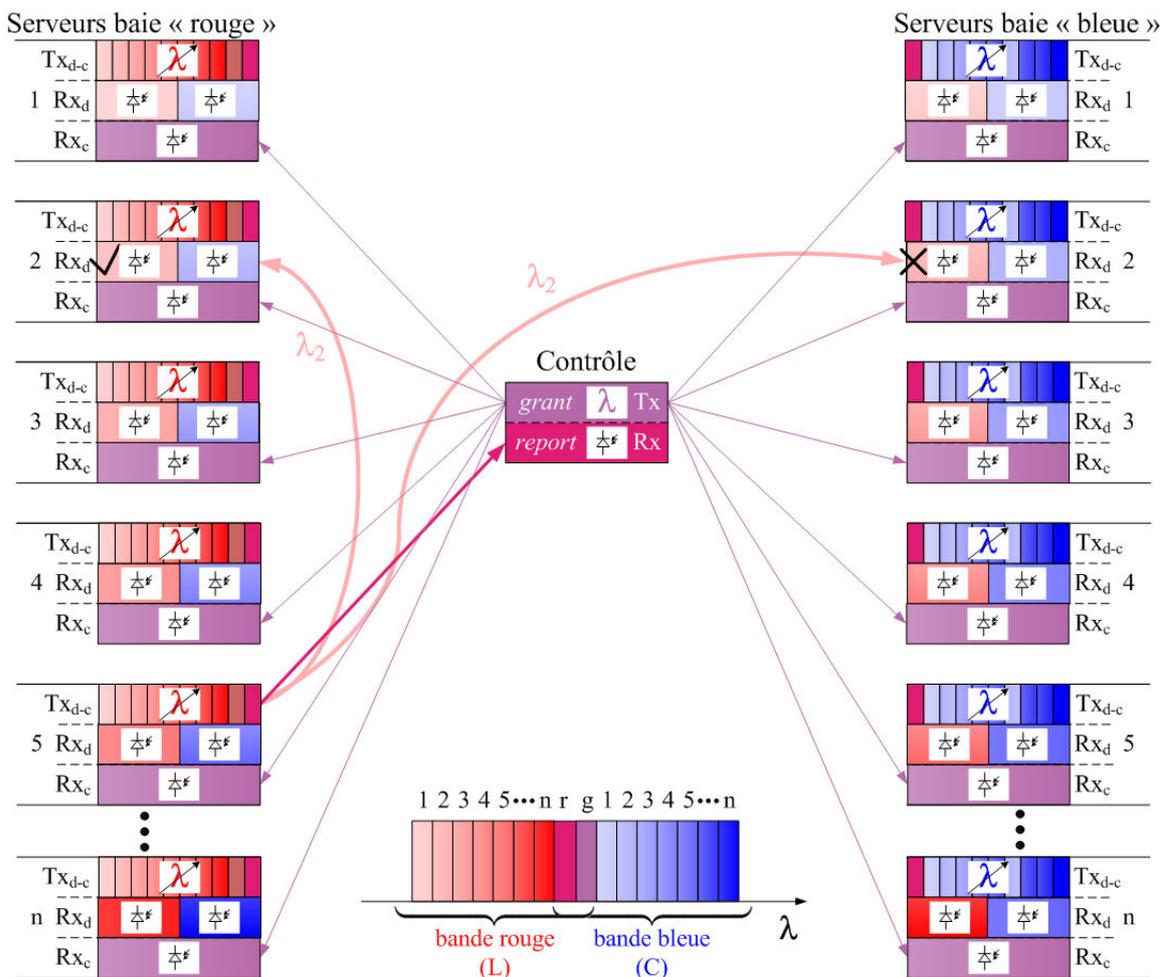
Chaque serveur est équipé d'un émetteur accordable en longueur d'onde avec une bande de transmission propre à sa baie et de deux récepteurs de données (chacun précédé par un filtre optique passe-bande pour la longueur d'onde relative à l'indicateur du serveur, un filtre pour chaque bande de transmission optique). Le schéma d'interconnexion de cette option est présenté sur la Figure 3.10(a); les structures de l'émetteur et du récepteur, sur la Figure 3.10(b).

La structure du contrôleur ne change pas, puisque nous utilisons toujours une seule paire de Tx/Rx pour l'envoi des messages de *report* et de *grant*. La complexité du plan de contrôle est accrue par rapport à POPI option D. Même si E-POPI 1 – 1 – 1 – 2 fonctionne avec deux récepteurs de données par serveur et suppose que ce dernier puisse traiter les deux flux provenant des bandes de transmission C ou L en même temps, cette option présente des possibles conflits qui doivent être gérés afin d'éviter les collisions.

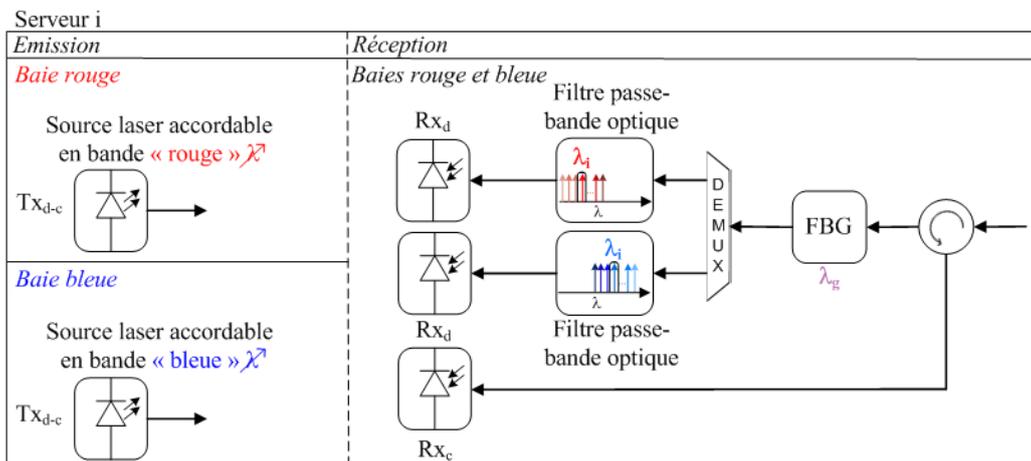
Un conflit similaire à celui observé dans la configuration 1 – 1 – 2 – 1 (Figure 3.9(b)) peut être observé sur la Figure 3.10(b). Le serveur 5 de la baie rouge envoie des paquets vers le serveur 2 de la même baie. Le serveur 2 de la baie bleue peut également recevoir ces paquets. Dans ce cas là nous n'aurions pas de conflit si un autre serveur de la baie bleue envoyait des paquets vers ce dernier, parce que ce serveur a deux récepteurs de données. Cependant, le conflit aurait lieu si un autre serveur de la baie rouge envoyait des paquets vers ce serveur. Le problème de la sécurité des données doit également être géré en ajoutant, par exemple, du cryptage. Le plan de contrôle de POPI original doit être adapté pour éviter ce genre de conflit.

La charge maximale autorisée n'est pas modifiée par rapport à POPI, parce que chaque serveur a un filtrage unique par récepteur de données. La possibilité de diffusion pourrait encore être exploitée en configurant le DC pour que les couples de serveurs les plus susceptibles de devoir s'en servir soient logés dans des baies différentes et sous le même indicateur de serveur. Le bilan de liaison et la consommation d'énergie varient par rapport à ceux de POPI. L'utilisation d'un démultiplexeur de bande en réception augmente les pertes du réseau de 1 dB. La consommation d'énergie de cette option augmente de 0,2 W par serveur, ceci étant dû au filtre optique passe-bande (bande C ou L) additionnel.

Nous ne considérons pas les structures 1–2–2–1 ou 1–2–1–2, parce qu'elles rajoutent de la complexité, augmentent le coût et la consommation d'énergie, sans augmenter la capacité, en comparaison avec les options précédentes.



(a)



(b)

Figure 3.10. E-POPI 1 – 1 – 1 – 2 : (a) schéma d’interconnexion, 1, 2, ..., n : longueurs d’onde relatives aux serveurs 1, 2, ..., n, respectivement, r et g : longueurs d’onde relatives aux messages d’*report* et de *grant*, respectivement ; (b) structure en émission et en réception.

Le Tableau 3.2 résume et compare les options d'E-POPI permettant de doubler le nombre de serveurs que nous venons de présenter en termes de capacité maximale potentielle, complexité du plan de contrôle, charge maximale autorisée, possibilité de diffusion, bilan de liaison et consommation d'énergie.

Critère	POPI	E-POPI		
	Option D	1 – 2 – 1 – 1	1 – 1 – 2 – 1	1 – 1 – 1 – 2
Capacité potentielle	61 serveurs	125 serveurs	125 serveurs	125 serveurs
Complexité du plan de contrôle	Référence	+	++	++
Charge maximale autorisée	Référence	=	- (-50%)	=
Possibilité de diffusion inter-baie	Non	Non	Oui	Oui
Augmentation de la consommation d'énergie	Référence	++ (3,5 W/serveur)	+ (+0,2 W/serveur)	+ (+0,2 W/serveur)

Légende : - diminution légère, = pas de variation, + augmentation légère, ++ augmentation modérée.

Tableau 3.2. Comparaison des différentes options d'E-POPI.

3.3.2 Options d'E-POPI à b bandes

Il s'agit des extensions des options 1 – 2 – 1 – 1, 1 – 1 – 2 – 1 et 1 – 1 – 1 – 2. La complexité du contrôleur est accrue parce que maintenant il utilise b paires Tx/Rx avec $b \geq 2$.

3.3.2.1 E-POPI b - b -1-1 ($b \geq 2$)

Chaque serveur est équipé de b émetteurs rapidement accordables en longueur d'onde sur b bandes de transmission optiques (le choix de l'émetteur correspond à la baie de destination) et d'un récepteur unique (un filtre optique en réception différent pour chaque serveur et chaque baie, un seul récepteur de données).

3.3.2.2 E-POPI b -1- b -1 ($b \geq 2$)

Chaque serveur est équipé d'un émetteur rapidement accordable en longueur d'onde avec une bande de transmission optique propre à sa baie et de b filtres optiques passe-bande en réception pour un seul récepteur de données. Ces filtres sont centrés à la longueur d'onde relative à l'indicateur du serveur. Le conflit présent dans l'option 1 – 1 – 2 – 1 est aussi observé dans $b – 1 – b – 1$, mais avec une complexité accrue, et doit être résolu de la même manière.

3.3.2.3 E-POPI $b-1-1-b$ ($b \geq 2$)

Chaque serveur est équipé d'un émetteur accordable en longueur d'onde avec une bande de transmission optique propre à sa baie et de b récepteurs de données (chacun précédé par un filtre optique passe-bande pour la longueur d'onde relative à l'indicateur du serveur). Les problèmes de collisions entre paquets son similaires à ceux de l'option 1 – 1 – 1 – 2.

3.4 Conclusion

Dans ce chapitre nous avons étudié la pertinence de POPI comme architecture d'interconnexion du projet EPOC. Pour ce faire, nous avons calculé la puissance d'émission minimale de POPI en fonction du nombre de machines connectées, en utilisant deux types de récepteur (PIN et APD). Nous avons effectué le même calcul sur POITR, celle-ci étant l'architecture basée sur OPS concurrente la plus intéressante en termes de consommation d'énergie. Les résultats ont montré que la puissance d'émission minimale est toujours moins élevée pour POPI et que les principales pertes de puissance de POPI sont dues aux pertes d'insertion du coupleur passif. De plus, nous avons proposé d'autres options de connexion des canaux de contrôle au coupleur central de POPI avec des pertes d'insertion très faibles.

Nous avons de même estimé le nombre maximal d'entités connectées, limité par l'émetteur accordable en longueur d'onde. Ce nombre est de 61 serveurs pour l'option D, en supposant des canaux WDM espacés de 50 GHz sur une bande de 3200 GHz. Dans ces conditions, avec des canaux au format de modulation NRZ le débit est limité à 25 Gbit/s. Nous étudierons des formats de modulation à plus grande efficacité spectrale dans le Chapitre 5.

Nous avons ensuite étudié la possibilité d'augmenter la capacité de POPI en termes de machines connectées limitées par les contraintes spectrales. Nous avons proposé E-POPI, extension de POPI qui nous permet de doubler le nombre d'entités connectées, en considérant deux groupes de serveurs utilisant des lasers accordables en longueur d'onde dans une bande de transmission propre au groupe concerné. Cette extension nous offre une possibilité de diffusion qui pourrait être potentiellement exploitée par la couche d'application du centre de données. De plus, E-POPI est extensible à plusieurs bandes de transmission.

Reste à explorer des possibilités d'augmenter le nombre de machines connectées aux réseaux POPI limités par les contraintes liées aux pertes d'insertion. Par exemple, en remplaçant une partie des coupleurs par d'autres composants ayant des pertes plus faibles ou en cascadeant deux ou plusieurs étages hiérarchiques. Nous étudions ces possibilités dans le Chapitre 4.

ARCHITECTURE D'INTERCONNEXION À DEUX ÉTAGES : POPI+

4.1 Introduction

Ce chapitre s'intéresse à augmenter d'une manière significative le nombre d'entités connectées à POPI. Une nouvelle architecture à deux étages basée sur TWIN est proposée. Nous revenons sur la notion de rack, chaque rack utilisant un réseau POPI indépendant. Les racks sont connectés grâce à une nouvelle structure, aussi basée sur TWIN, utilisant des WSS. Nous décrivons la conception de cette nouvelle structure inter-racks et les éventuelles adaptations de POPI qui doivent être prises en compte sur les connexions intra-racks.

Nous étudions ensuite le bilan de liaison de cette architecture. Ce calcul nous permettra d'estimer le nombre maximal de serveurs connectés dans le chapitre suivant.

4.2 Description de l'architecture POPI+

Nous introduisons une structure inter-racks, également basée sur TWIN. Le principe de ce système est de connecter l'émetteur accordable en longueur d'onde à un coupleur passif $N : 1$, dont la sortie commune est reliée à un démultiplexeur $1 : N$, dont chacune de ses sorties est connectée à un récepteur optique, selon la Figure 4.1.

Cette figure représente le schéma d'interconnexion basée sur le routage de longueur d'onde le plus basique permettant d'appliquer le principe de TWIN (sachant que POPI constitue une version “*broadcast & select*” de TWIN). Dans cette architecture, aussi bien que sur POPI, les intervalles de temps (*time slots*) peuvent être attribués grâce à un algorithme d'allocation de bande passante basé sur PON, qui peut-être dynamique ou statique selon la variabilité du trafic.

La structure « coupleur + démultiplexeur » de la Figure 4.1 peut être implémentée en utilisant des composants divers. Par exemple, le coupleur $N : 1$ pourrait être remplacé par une cascade de coupleurs de moindre capacité (selon la disponibilité sur le marché).

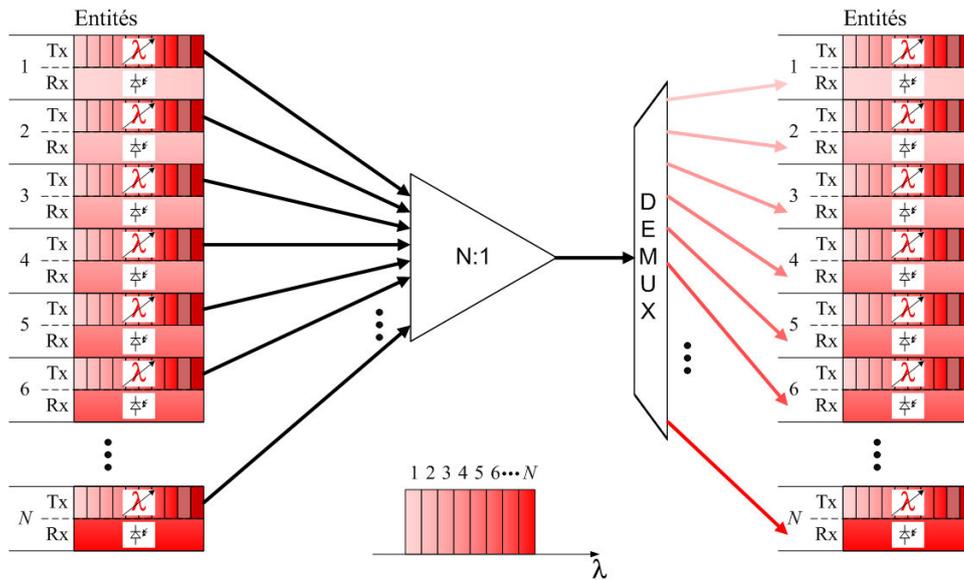


Figure 4.1. Architecture d'interconnexion TWIN basique basée sur un coupleur $N : 1$ et un démultiplexeur (DEMUX) $1 : N$. $1, 2, \dots, N$: longueurs d'ondes relatives aux entités (racks ou serveurs) $1, 2, \dots, N$, respectivement.

Une autre possibilité serait de remplacer le démultiplexeur par des WSS, car ces derniers peuvent présenter plusieurs avantages pour les DC. Parmi ces fonctionnalités : des pertes qui se comparent favorablement avec celles du démultiplexeur (≤ 5 dB pour des WSS Finisar ou JDSU [113]-[115]) ; une haute résolution spectrale, développée dans le cadre des systèmes de transmission WDM FlexGrid, permettant des pas aussi bas que 12,5 GHz ; et un profil spectral généralement plus rectangulaire que celui du démultiplexeur (même comparé à des démultiplexeurs *flat top*, dont les pertes sont bien plus élevées [116]). De plus, la possibilité de pouvoir changer de temps en temps le routage spectral donne plus de flexibilité opérationnelle au réseau.

A présent, le principal inconvénient des WSS est le nombre limité de ports de sortie, qui empêche leur utilisation comme démultiplexeur reconfigurable pour un grand nombre de canaux, par exemple, pour la bande de transmission C au pas de 50 GHz. Des WSS à 4, 9 et 20 ports seulement sont disponibles dans le commerce actuellement. Néanmoins, sachant que les pertes d'insertion des WSS ne dépendent que légèrement du nombre de ports de sortie (≤ 5 dB pour 20 ports [115]), une cascade de deux WSS, serait déjà suffisante pour émuler un démultiplexeur reconfigurable FlexGrid, présentant des pertes inférieures à 10 dB.

D'un autre côté, la structure proposée doit inclure un moyen de compenser les pertes du coupleur passif $N : 1$, lorsque nous augmentons le nombre d'entités connectées. Ceci peut être résolu en ajoutant un amplificateur optique entre le coupleur et la cascade de WSS.

Dans le contexte des DC, une structure TWIN basée sur des WSS pourrait en principe être utilisée pour desservir les connexions intra-racks aussi bien que les inter-racks. Toutefois, nous considérons que l'utilisation de cette structure à l'intérieur de chaque rack générerait

des coûts et consommation d'énergie additionnels très élevés (composants actifs : WSS¹ et amplificateur optique²) par rapport à POPI (composants purement passifs). Alors que, ces coûts supplémentaires sont moins pénalisants dans le cadre d'une structure de connexion commune entre les racks.

En somme, nous nous intéressons à une architecture OPS à deux étages :

1^{er} étage : connexions intra-racks basées sur POPI.

2^{ème} étage : connexions inter-racks assurées par une structure basée sur des WSS.

La Figure 4.2 présente la nouvelle structure inter-racks proposée. Chaque rack possède une interface constituée par un émetteur accordable en longueur d'onde et un récepteur à une longueur d'onde fixe. Chaque rack est connecté à un coupleur passif $N + 1 : 1$ (où N représente le nombre de racks), suivi par un amplificateur optique pour compenser les pertes du coupleur et une cascade de WSS $1 : 9$ à deux étages.

$N + 1$ pourrait s'élever jusqu'à 81 (racks et *gateway* du DC), avec des WSS $1 : 9$. Cependant, ce chiffre sera au moins réduit à 64, capacité de l'émetteur accordable en longueur d'onde [26]-[27] (lasers identiques à ceux considérés pour POPI).

En outre, le choix de l'amplificateur optique prend en compte le fait qu'un rack de destination ne peut pas recevoir des paquets/rafales provenant de plusieurs sources en même temps. Autrement dit, un unique serveur parmi la totalité des serveurs du DC (hébergé par n'importe quel rack) ne peut émettre que vers un rack de destination (différent du sien) à la fois. De ce fait, l'amplificateur optique aura un maximum de 64 longueurs d'ondes à amplifier, correspondant à la limite du laser accordable. La charge de cet amplificateur peut varier donc de 0 à 64 canaux à une vitesse de l'ordre de la dizaine de nanoseconde. L'amplificateur à fibre dopée à l'erbium (*EDFA : Erbium-Doped-Fiber Amplifier*) est potentiellement un bon candidat à cet effet pour les raisons suivantes :

- Il présente une dynamique lente (≈ 10 ms [119]) par rapport au temps d'accordabilité en longueur d'onde de l'émetteur (≤ 50 ns [26]-[27]), ce qui le rend insensible à la présence éventuelle de temps de garde et aux changements de longueurs d'onde entre paquets successifs.
- En cas de variations sur une longue période de la charge de l'EDFA (soit en nombre soit en position de canaux), plusieurs techniques de contrôle sur la variation de gain ont été étudiées et implémentées. Parmi ces techniques : contrôle tout-optique ou électronique-optique grâce à l'introduction d'un canal de compensation [120]-[121] ou en modifiant la puissance de la pompe de l'amplificateur [122].
- Disponibilité commerciale dans les bandes de transmission optiques C et L, ce qui permet l'utilisation d'E-POPI [117]-[118].

1. environ 1000 USD et 4 W par port (soit 10.000 et 400 W pour deux étages de WSS $1 : 9$) [76], [115].

2. environ 10.000 ou 12.000 USD pour un EDFA en bande de transmission C ou L, respectivement ; ≤ 12 W pour un gain de 20 dB (puissance d'entrée ≈ -10 dBm) [117]-[118].

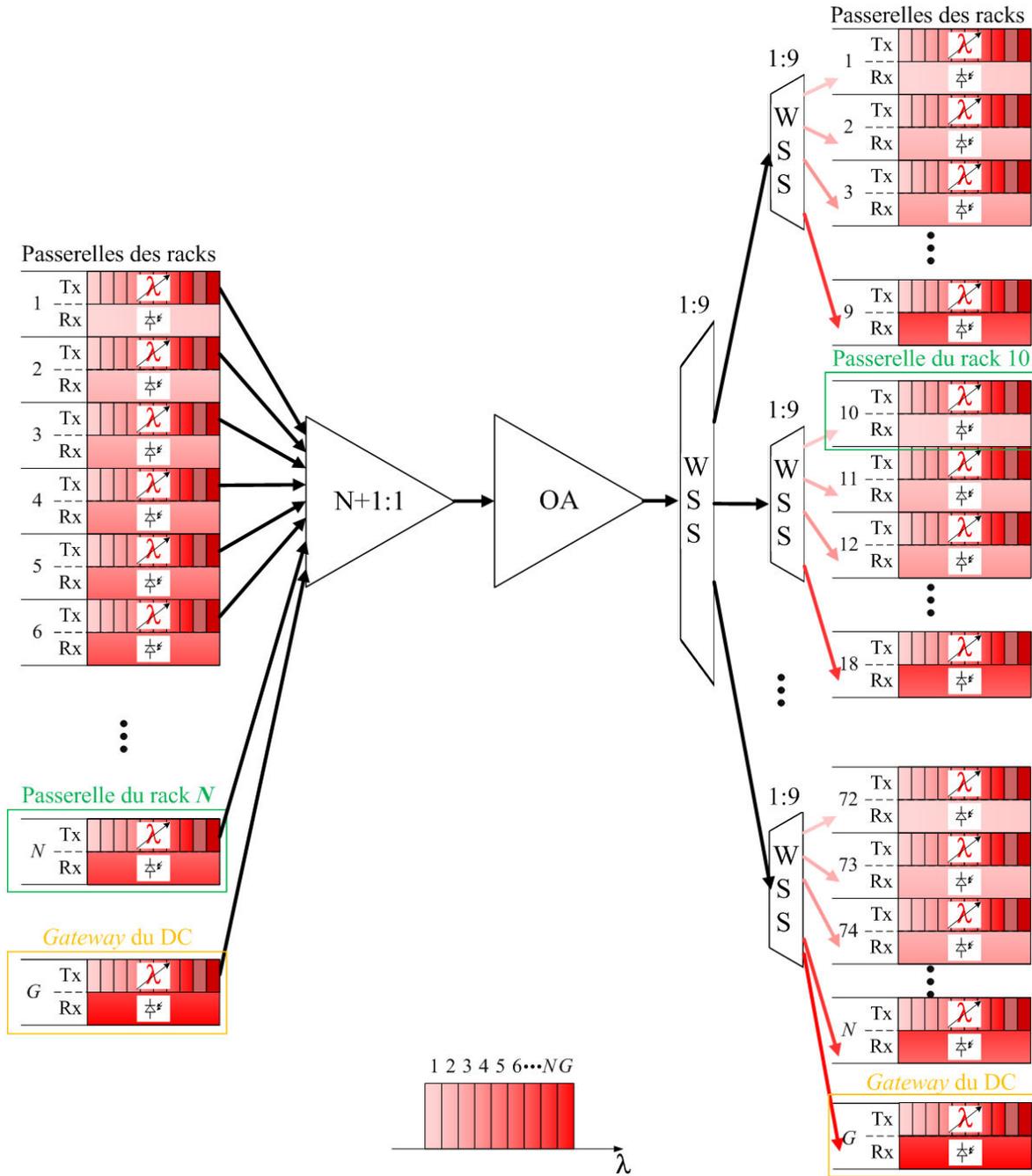


Figure 4.2. Architecture d'interconnexion TWIN adaptée aux connexions inter-racks. Structure basée sur un coupleur $N + 1 : 1$, un amplificateur optique (OA) et deux étages de WSS $1 : N$. 1, 2, ..., N, G : longueurs d'ondes relatives aux racks 1, 2, ..., N et au gateway du DC, respectivement.

— Disponibilité commerciale à des puissances de saturation très élevées (jusqu'à 23 dBm) [117]-[118], permettant d'augmenter la puissance d'émission, et donc le nombre de racks connectés voire le nombre d'états de modulation.

L'idée générale, inspirée d'E-POPI (1 – 2 – 1 – 1), consiste à utiliser deux émetteurs

accordables en longueur d'onde par serveur, opérant sur les bandes de transmission optiques C et L, respectivement. Une des ces deux bandes de transmission (par exemple la bande C) sera destinée aux connexions intra-racks ; et l'autre (par exemple la bande L), aux connexions inter-racks. Le fait de considérer une extension hiérarchique sur POPI+, permet d'augmenter considérablement le nombre de serveurs connectés par rapport à des solutions à un étage. POPI+ peut connecter jusqu'à $N \times n$ serveurs, N et n étant le nombre maximal de racks et de serveurs par rack connectés au réseau POPI standard (option D), respectivement.

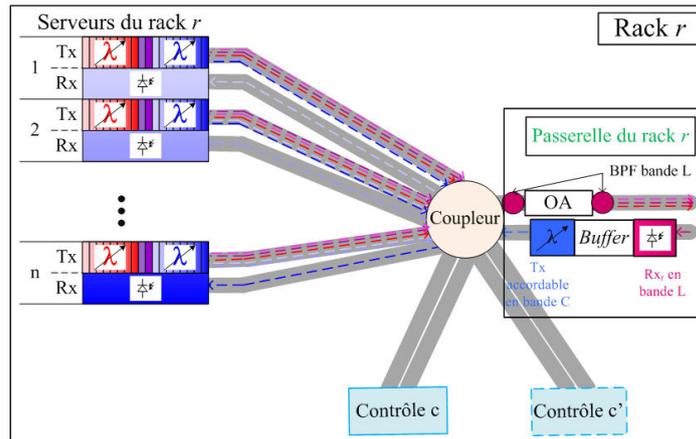
Le réseau POPI intra-rack doit être adapté pour ainsi être connecté à la structure inter-racks. Nous présentons le schéma d'interconnexion de POPI intra-rack, les structures en émission et en réception des serveurs et de la passerelle sur les figures 4.3(a), (b) et (c), respectivement.

Les serveurs logés dans un rack donné vont utiliser le réseau POPI en bande de transmission C pour les connexions intra-racks. L'émetteur accordable en bande C servira pour la transmission des données à l'intérieur du rack et du *report* intra-rack qui sera reçu par les contrôleurs locaux de POPI. Les racks seront connectés par une deuxième interface (la structure basée sur un coupleur, un OA et deux étages de WSS que nous venons d'introduire) opérant dans la bande de transmission optique L. Ainsi l'émetteur accordable en bande L servira pour la transmission des données vers la structure inter-racks et du *report* inter-racks qui sera reçu par une unité de contrôle centrale commune à tous les racks du DC. En ce qui concerne les connexions intra-racks, le bilan de liaison d'E-POPI $1 - 2 - 1 - 1$ est toujours valable, car la structure des serveurs reste la même.

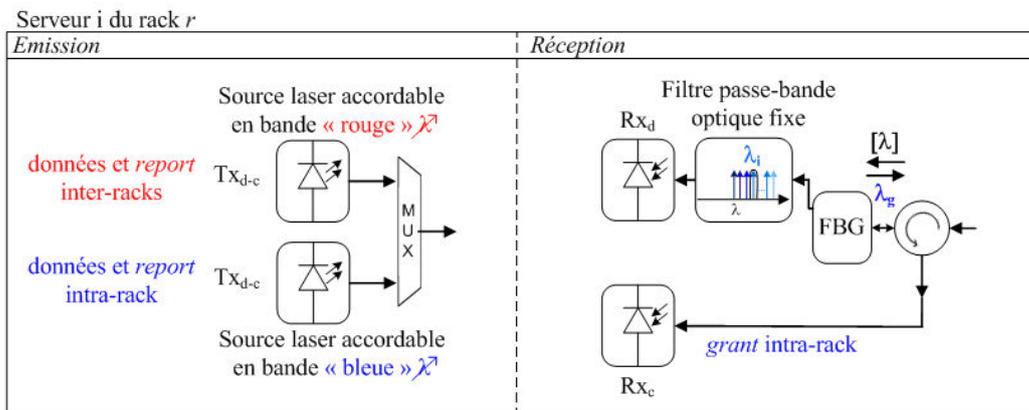
Autrement dit, un serveur qui aurait besoin de se connecter avec un autre appartenant à un rack différent, va utiliser l'émetteur accordable de la bande L et ajuster sa longueur d'onde en accord avec l'indicateur du rack dans lequel le serveur destinataire est logé. Ces paquets vont accéder au réseau POPI local de ce rack grâce à la « passerelle du rack » (Figure 4.3(c)). Le flux montant ou *upstream* (se dirigeant vers l'extérieur du rack) utilise une interface complètement transparente. Elle consiste en un amplificateur optique (EDFA) et un filtre passe-bande optique en bande L, destinés à compenser les pertes d'insertion du coupleur et bloquer le trafic en bande C, respectivement. Le flux descendant ou *downstream* (en provenance de l'extérieur du rack) utilise une interface opaque. Les paquets arrivent de manière complètement transparente à la passerelle du rack, où ils vont être reçus, traités et régénérés grâce à un émetteur accordable en longueur d'onde de la bande C. Ce dernier ajustera sa longueur d'onde en fonction du serveur de destination.

Afin de mieux comprendre le fonctionnement global de l'architecture proposée, la Figure 4.4 présente le schéma d'interconnexion complet de POPI+, avec les interfaces intra et inter-racks.

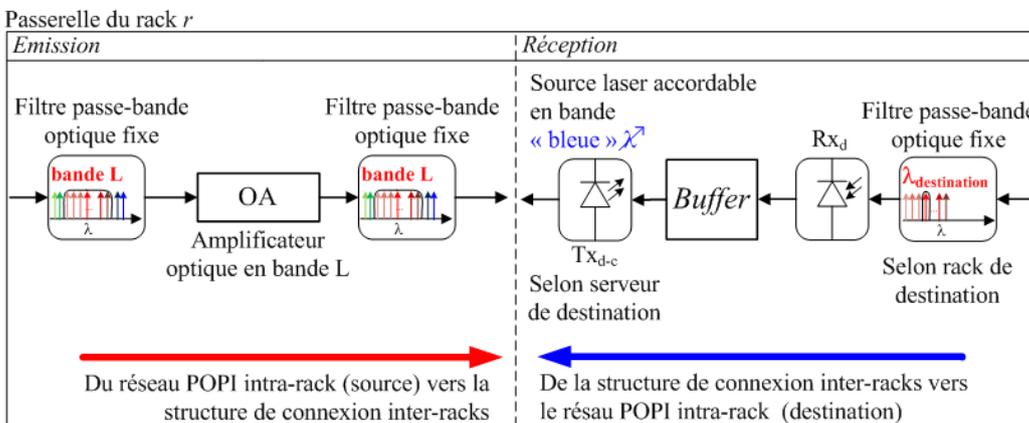
Une unité centrale de contrôle va gérer les connexions inter-racks grâce à un réseau indépendant de la structure inter-racks, nous présentons ce réseau sur la Figure 4.5. Il est basé sur un coupleur passif (comme celui du POPI) $N + 1 : N + 1$ qui va connecter l'unité



(a)



(b)



(c)

Figure 4.3. POPI adaptée pour les connexions intra-racks du réseau à deux étages POPI+. (a) Schéma d'interconnexion. Connexions intra-racks en bande de transmission optique C (bleue); et inter-racks en bande de transmission L (rouge). Par simplicité nous avons omis les récepteurs de contrôle (toujours présents dans l'architecture). (b) Structure des serveurs en émission et en réception. (c) Structure de la passerelle du rack en émission et en réception.

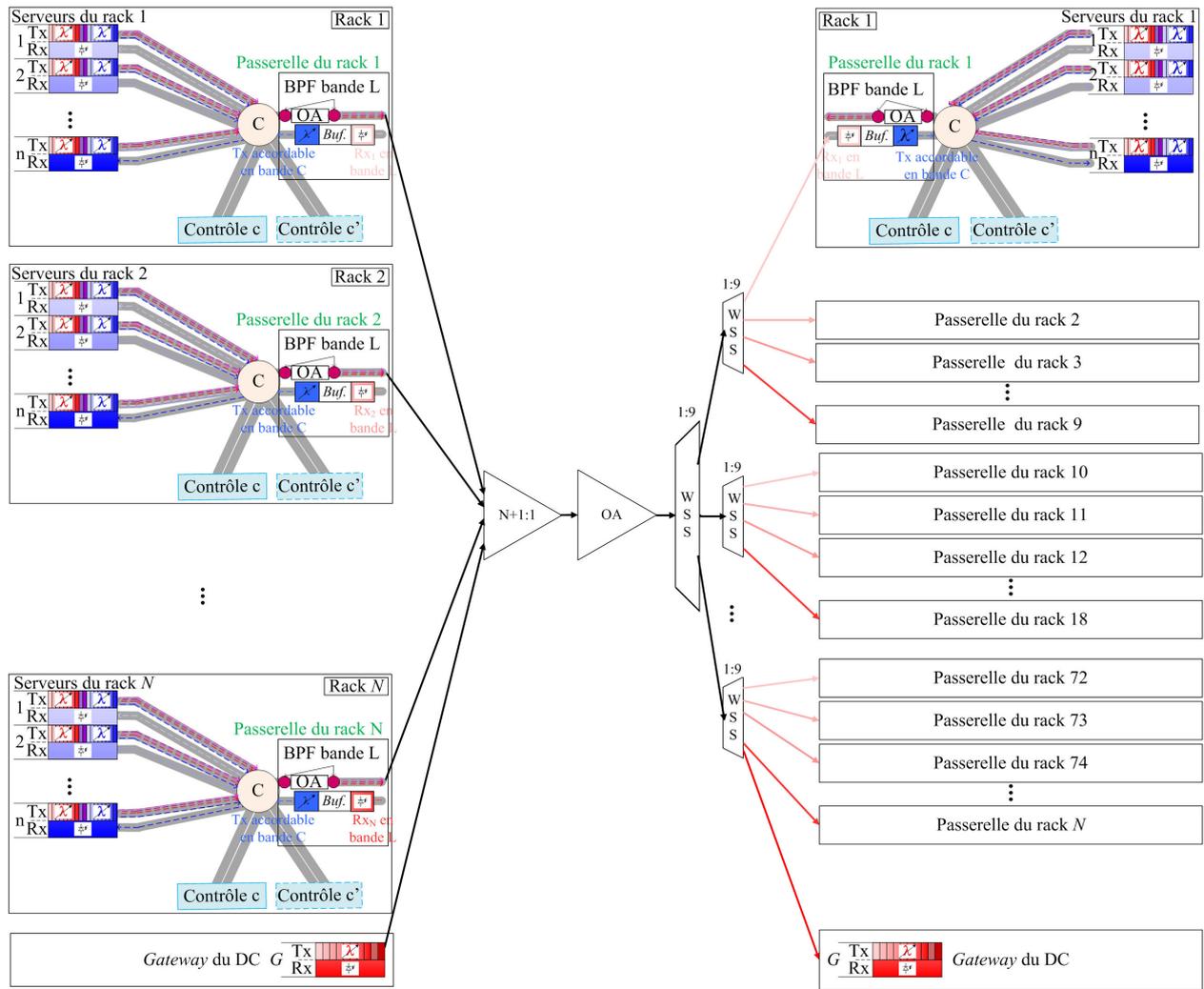


Figure 4.4. Schéma d'interconnexion complet POPI+. Serveurs source (colonne de gauche) et serveurs de destination (colonne de droite). C : Coupleur ; Buf. : Buffer.

centrale de contrôle à tous les contrôleurs locaux de chaque rack. L'unité centrale de contrôle reçoit des *reports* en bande de transmission L provenant de tous les serveurs du DC (λ_r); ainsi, elle délivre des *grants* aux contrôleurs locaux (λ_g)³. Ces derniers ont deux récepteurs : le premier va recevoir les *reports* en bande de transmission C pour les connexions intra-racks (λ_r); le second, va recevoir les *grants* inter-racks issus par l'unité centrale de contrôle (λ_g). Les contrôleurs locaux ont un unique émetteur en longueur d'onde fixe (λ_g) qui sera utilisé pour envoyer des messages de *grant* aux serveurs pour les connexions intra et inter-racks.

Notons que les communications depuis les serveurs vers l'unité centrale de contrôle (λ_r) passent par une interface transparente. De ce fait il est nécessaire d'utiliser un amplificateur optique (SOA ou EDFA) après passage à travers le coupleur de POPI afin de pouvoir compenser ses pertes d'insertion et de pouvoir cascader les deux réseaux. Par contre, l'interface

3. Peu importe la bande de transmission optique ou la longueur d'onde du *grant* inter-racks, il faut juste qu'elle soit différente de la longueur d'onde du *report* inter-racks (λ_r).

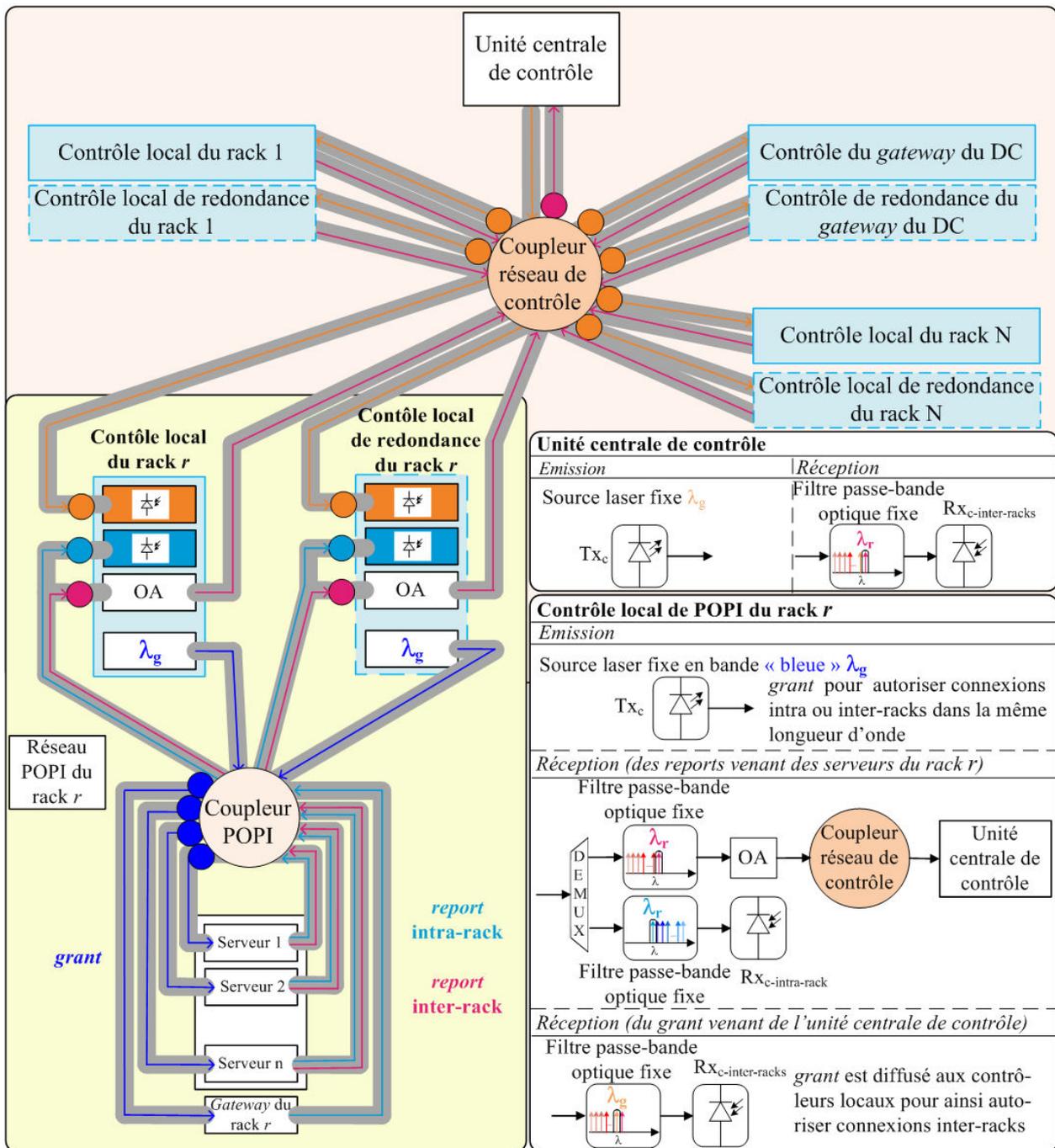


Figure 4.5. Schéma d'interconnexion du réseau de contrôle de POPI+.

depuis l'unité de centrale de contrôle vers les serveurs est opaque. Ceci est dû au fait que les autorisations d'envoi de messages (λ_g) sont générées par contrôleur local sur la base des informations de disponibilité de la structure inter-racks (λ_g) et également de celles du réseau POPI intra-rack.

La structure des paquets (ou rafales) change légèrement selon la nature de la connexion : intra ou inter-racks, car pour ces dernières, l'information sur le rack de destination (donnée

par la longueur d'onde en bande L) n'est pas suffisante, il faut encore pouvoir identifier le serveur de destination une fois les paquets arrivés à la passerelle du rack de destination. Cette information peut être incorporée dans l'entête des paquets sous forme d'adresse de destination ou label.

En conséquence, les connexions inter-racks vont devoir utiliser, en première instance, l'unité centrale de contrôle qui communiquera aux contrôleurs locaux la disponibilité la structure inter-racks. Les contrôleurs locaux ayant déjà la disponibilité du réseau intra-rack, vont pouvoir ainsi délivrer l'autorisation pour accéder à la passerelle du rack source et à la structure inter-racks. Finalement, une fois le paquet dans la mémoire tampon au niveau de la passerelle du rack de destination (avant d'être régénéré en bande C), une nouvelle demande de connexion sera envoyée aux contrôleurs locaux de ce rack. Une fois délivrée cette autorisation, le paquet sera régénéré par l'émetteur accordable de cette passerelle à la longueur d'onde du serveur de destination.

Le fait de devoir partager les contrôleurs locaux de POPI avec celui des connexions inter-racks ne pose pas de problème majeur, car dans un DC de petite/moyenne taille les transferts intra-racks sont beaucoup plus fréquents que les inter-racks (environ 80% contre 20%) [96]-[97].

En outre, nous estimons la capacité maximale potentielle de POPI+. Pour cela nous prenons en compte le fait que le réseau POPI intra-rack doit réserver deux longueurs d'onde de la bande de transmission C pour le *report* (λ_r) et le *grant* (λ_g); ainsi que la structure inter-racks doit réserver trois longueurs d'onde de la bande de transmission L pour le *report* (λ_r), le *grant* (λ_g) et le *gateway* du DC (λ_G). De sorte que, la capacité maximale potentielle de POPI+, en considérant l'émetteur accordable présenté dans [26]-[27], est de 61 racks de 62 serveurs chacun.

4.2.1 Variante de POPI+ : Transfert-POPI

Dans cette section nous décrivons Transfert-POPI (T-POPI), réseau inspiré de POPI+, qui est également basé sur le principe E-POPI en deux étages. T-POPI facilite les transferts de données massifs. Cette option consiste à placer un deuxième récepteur de données à chaque serveur, qui serait dédié exclusivement aux paquets transmis en bande L. C'est-à-dire, chaque serveur aurait deux interfaces de données (et de contrôle) dédiées chacune à un étage du réseau :

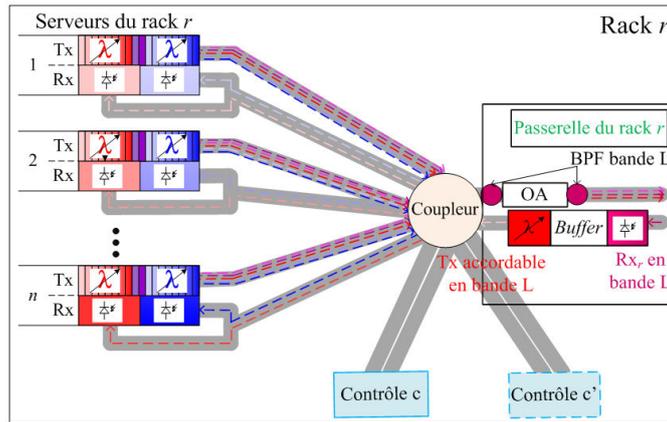
- **Réseau intra-rack basé sur POPI** : chaque serveur a un émetteur accordable en longueur d'onde en bande C (longueur d'onde ajustée en accord avec l'indicateur du serveur de destination) et un récepteur à longueur d'onde fixe en bande C (en accord avec l'indicateur du serveur).
- **Réseau inter-racks basé sur des WSS** : chaque serveur a un émetteur accordable en longueur d'onde en bande L (longueur d'onde ajustée en accord avec l'indicateur du rack de destination) et un récepteur à longueur d'onde fixe en bande L (en accord avec l'indicateur du serveur).

T-POPI offre des connexions inter-racks exclusives, qui ne doivent pas partager le réseau avec le trafic interne du rack. Cette fonctionnalité pourrait être fortement exploitée par des DC utilisant, par exemple, la politique VOVO décrite dans le Chapitre 2.

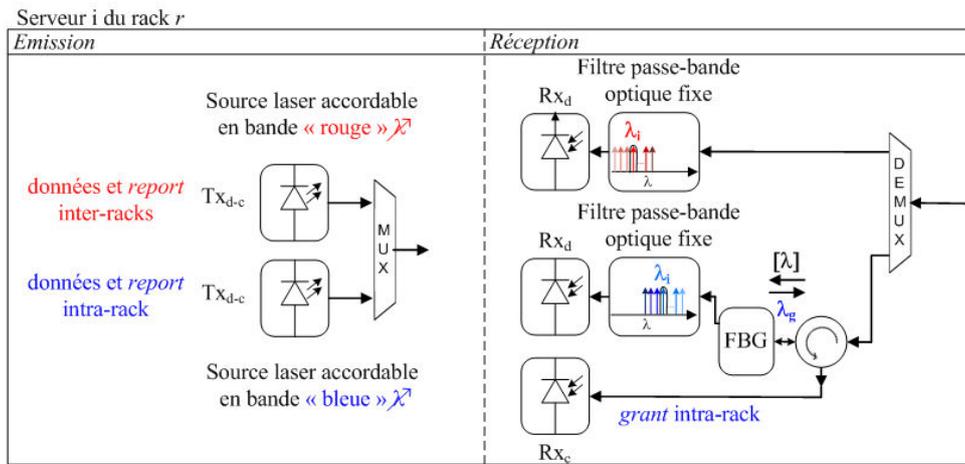
Nous présentons le schéma d'interconnexion de T-POPI (étage intra-rack), les structures en émission et en réception des serveurs et de la passerelle sur les figures 4.6(a), (b) et (c), respectivement.

La passerelle du rack de destination de T-POPI, après avoir reçu des paquets provenant d'autres racks ou du *gateway* du DC (jusqu'ici rien n'a changé par rapport à POPI+), va faire l'adressage des paquets vers le serveur de destination grâce une régénération en bande L (au lieu de bande C), voir figures 4.6(a) et (c). L'interface des serveurs en émission ne change pas par rapport à POPI+ ; cependant, celle du récepteur va inclure une structure de réception de données et de contrôle pour la bande de transmission L, un démultiplexeur de bandes C et L permettra de séparer le flux et l'adresser la structure de réception en bande C ou L, voir Figure 4.6(a). Ce dernier augmente les pertes du bilan de liaison intra-rack d'1 dB par rapport à POPI+.

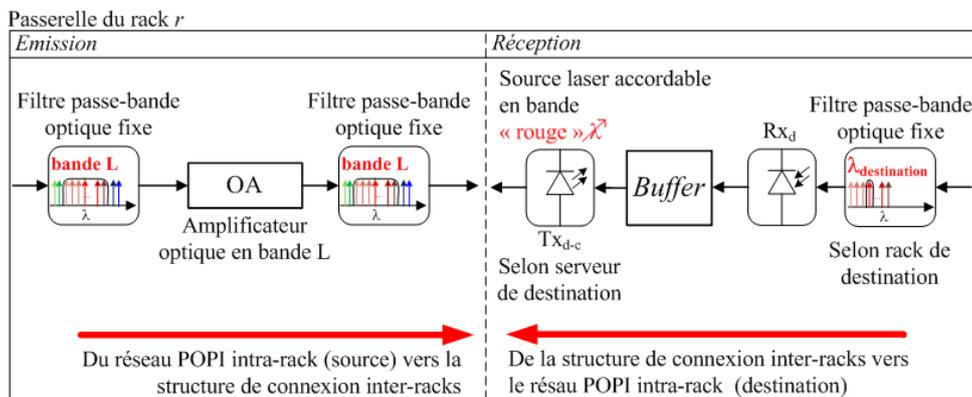
Le fait de régénérer le flux descendant en bande L évite toute sorte de conflit avec les connexions intra-racks en bande C. De plus, il ne peut y avoir de conflit entre les différents canaux de la bande L, puisqu'il ne peut y avoir qu'un seul flux descendant à la fois. C'est-à-dire, de ce point vue nous n'aurions pas besoin de faire une demande de connexion aux contrôleurs locaux du rack de destination. Cependant, ceci supposerait que chaque serveur peut traiter les flux intra et inter-racks en même temps. Nonobstant, si ce n'est pas le cas, un message de *report* doit être envoyé avant la régénération du signal. Nous modifions en conséquence la structure en réception des contrôleurs locaux, afin de pouvoir gérer trois messages de *report* différents : 1. intra-rack (λ_r) ; 2. inter-racks pour accéder à la passerelle du rack source et à la structure inter-racks (λ_r) ; et inter-racks pour accéder au réseau POPI du rack de destination (λ_r). La structure des contrôleurs locaux que nous proposons est présenté sur la Figure 4.7.



(a)



(b)



(c)

Figure 4.6. POPI adaptée pour les connexions intra-racks du réseau à deux étages T-POPI. (a) Schéma d'interconnexion. Connexions intra-racks en bande de transmission optique L (rouge); et inter-racks également en bande de transmission L (rouge). Par simplicité nous avons omis les récepteurs de contrôle (toujours présents dans l'architecture). (b) Structure des serveurs en émission et en réception. (c) Structure de la passerelle du rack en émission et en réception.

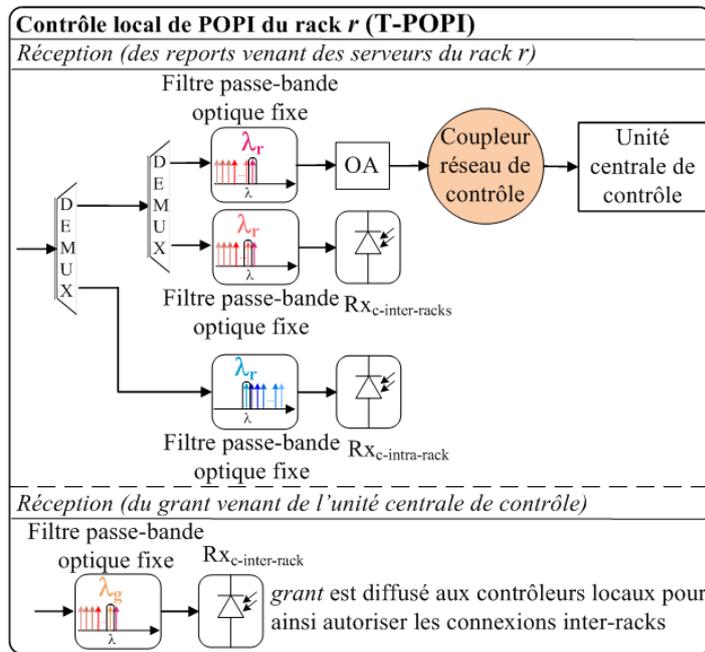


Figure 4.7. Structure des contrôleurs locaux en émission et en réception.

Nous estimons de même la capacité maximale potentielle de T-POPI. Nous tenons compte du fait que le réseau POPI intra-rack doit toujours réserver deux longueurs d'onde de la bande de transmission C pour le *report* (λ_r) et le *grant* (λ_g); et que la structure inter-racks doit réserver quatre longueurs d'onde de la bande de transmission L pour les deux messages de *report* (λ_r et λ_r), le message de *grant* (λ_g) et le *gateway* du DC (λ_G). Ainsi, la capacité maximale potentielle de T-POPI, en considérant le même émetteur accordable, est de 60 racks de 62 serveurs chacun.

4.2.1.1 Variante de T-POPI

Nous explorons maintenant une variante de T-POPI qui n'aurait pas besoin de l'utilisation de :

- multiplexage WDM;
- mise en mémoire tampon;
- régénération du signal.

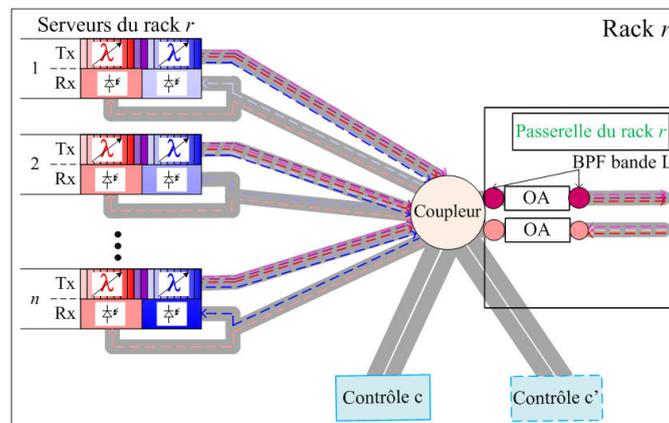
Il s'agit d'une option dans laquelle les interfaces inter-racks sont complètement transparentes dans les directions montante et descendante.

D'une part, nous partons du principe qu'un seul flux peut être envoyé vers un rack de destination (direction descendante) à la fois. Ceci nous permettrait de remplacer l'émetteur accordable en bande de transmission L de la passerelle de T-POPI par un émetteur en longueur d'onde fixe et de détecter le serveur de destination au niveau de chaque serveur

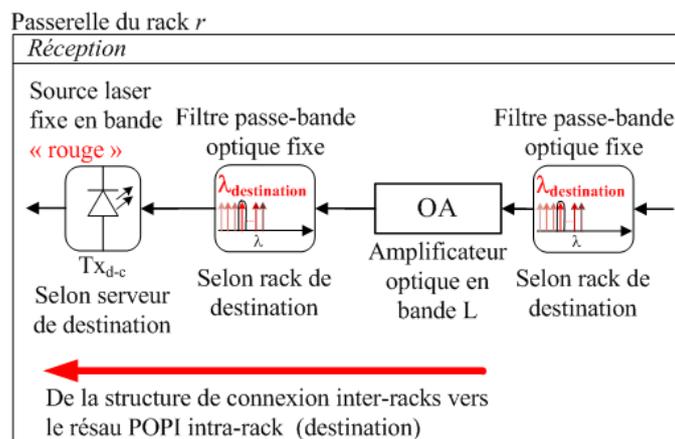
grâce à l'entête des paquets.

D'autre part, il ne peut pas y avoir de conflit entre la longueur d'onde régénérée en bande L et celles des connexions intra-racks en bande C. Nous n'avons donc pas besoin de faire de demande de connexion aux contrôleurs locaux du rack de destination⁴. En conséquence, nous n'aurions pas besoin de mettre en mémoire tampon ni de régénérer le flux descendant. Cependant, il serait nécessaire d'ajouter un amplificateur optique dans la direction descendante afin de compenser les pertes du réseau POPI intra-rack du rack de destination. Ceci pourrait être éventuellement contraignant, car la sensibilité des récepteurs serait dégradée.

Nous présentons le schéma d'interconnexion et la structure de la passerelle du rack en réception de cette variante de T-POPI sur les figures 4.8(a) et (b), respectivement (les structures des serveurs en émission et en réception, ainsi que celle du *gateway* en émission ne changent pas par rapport à T-POPI).



(a)



(b)

Figure 4.8. Variante de T-POPI : (a) schéma d'interconnexion intra-rack ; (b) structure en réception de la passerelle du rack.

4. Il faudrait que les contrôleurs locaux de chaque rack et l'unité centrale de contrôle soient au courant de l'état de la mise en mémoire tampon des flux intra et inter-racks au niveau de chaque serveur du DC.

Cette version de T-POPI est moins chère que l'option originale, mais elle pourrait présenter des éventuels soucis de sécurité concernant la diffusion des paquets à tous les serveurs du rack pour les connexions inter-racks. La mise en place d'un mécanisme de cryptage pourrait protéger ces données. Ce mécanisme serait appliqué en fonction de l'entête du paquet, de façon similaire à l'opération réalisée pour les flux descendants dans un réseau PON.

Finalement, nous estimons la capacité maximale potentielle de la variante de T-POPI. Pour le réseau POPI intra-rack deux longueurs d'onde de la bande de transmission C pour le *report* (λ_r) et le *grant* (λ_g) doivent toujours être réservées. Pour la structure inter-racks trois longueurs d'onde de la bande de transmission L pour le *report* (λ_r), le *grant* (λ_g) et le *gateway* du DC (λ_G) doivent être également réservées. La capacité maximale potentielle de la variante de T-POPI est donc de 61 racks de 62 serveurs chacun.

Le Tableau 4.1 résume et compare POPI+, T-POPI et sa variante en termes de capacité maximale potentielle, complexité du plan de contrôle, charge maximale autorisée, possibilité de diffusion, bilan de liaison et consommation d'énergie.

Critère	POPI Option D	POPI+	T-POPI	Variante de T-POPI
Capacité potentielle	61 serveurs	3782 serveurs	3720 serveurs	3782 serveurs
Complexité du plan de contrôle	Référence	+	+	++
Charge maximale autorisée	Référence	=	=	=
Possibilité de diffusion inter-baie	Non	Non	Non	Non
Augmentation des pertes dans le bilan de liaison intra-rack	Référence	+ (+2 dB)	++ (+3 dB)	++ (+3 dB)
Augmentation de la consommation d'énergie (étage intra-rack)	Référence	++ (+3,7 W/serveur)	+ (+3,7 W/serveur)	+ (+1,2 W/serveur)

Légende : - diminution légère, = pas de variation, + augmentation légère, ++ augmentation modérée.

Tableau 4.1. Comparaison des options POPI+, T-POPI et variante de T-POPI.

4.3 Bilan de liaison et dimensionnement de POPI+

Dans cette section nous présentons le bilan de liaison de POPI+ et nous estimons sa capacité potentielle. Dans un premier temps nous estimons le bilan de liaison intra-rack de POPI+ grâce à l'équation 4.3.1, basée sur les bilans de liaison de POPI (équation 3.2.3 +1 dB de pertes d'insertion liées au circulateur ($L_{circulateur}$) et au FBG (L_{FBG}) de l'option D), et celui d' E-POPI 1 – 2 – 1 – 1 (plus 1 dB de pertes d'insertion liées au multiplexeur

de bandes C et L (L_{mux}). Ce bilan nous permettra de déterminer le nombre maximal de serveurs connectés au réseau intra-rack.

$$\begin{aligned} \check{P}_{Tx,POPI\ intra-rack} \geq & P_{Rx,intra-rack} + L_{fibre,intra-rack} + L_{connecteurs} \\ & + L_{EAM} + L_{BPF,intra-rack} + L_{C(total),intra-rack} \\ & + L_{mux} + L_{circulateur} + L_{FBG} + M_{intra-rack} \end{aligned} \quad (4.3.1)$$

Où $\check{P}_{Tx,POPI\ intra-rack}$ représente la puissance d'émission minimale de la structure intra-rack.

Ensuite, estimons le bilan de liaison du réseau inter-racks à l'aide de l'équation 4.3.2.

$$\begin{aligned} \check{P}_{in,inter-racks} \geq & P_{Rx,inter-racks} + L_{fibre,inter-racks} + L_{connecteurs} \\ & + L_{C(total),inter-racks} - G_{OA(in-line)} \\ & + 2 \times L_{WSS} + M_{inter-racks} \end{aligned} \quad (4.3.2)$$

Où $\check{P}_{in,inter-racks}$ correspond à la puissance minimale en entrée de la structure inter-racks basée sur des WSS (passerelle du rack en direction ascendante). $P_{Rx,inter-racks}$ est la puissance reçue requise pour assurer un BER de 10^{-12} (passerelle du rack en direction descendante). Ici nous considérons $P_{Rx,inter-racks}$ égal à la sensibilité des récepteurs PIN et APD du Tableau 3.1 plus les pénalités de puissance dues aux deux amplificateurs optiques utilisés sur l'étage inter-racks ($\leq 0,5$ dB, pour des débits inférieurs à 10 Gbit/s [123]). $L_{fibre,inter-racks}$ et $L_{connecteurs}$ sont les pertes dues à la fibre optique (1 km) et aux connecteurs, respectivement. $L_{C(total),inter-racks}$ représente les pertes du coupleur $N + 1 : 1$, estimées grâce à l'équation 3.2.6. $G_{OA(in-line)}$ correspond au gain de l'OA placé entre le coupleur $N + 1 : 1$ et la cascade de WSS. L_{WSS} représente les pertes de chaque WSS. $M_{inter-racks}$ est une marge de 1,5 dB.

Le Tableau 4.2 montre les valeurs de sensibilité des récepteurs (PIN et APD), les pertes et les gains des composants pris en compte afin de pouvoir estimer $\check{P}_{in,inter-racks}$.

Composant	$P_{Rx,inter-racks}$ (dBm)	Référence
PIN	$-19 + 0,5 = -18,5$	[83], [123]
APD	$-24 + 0,5 = -23,5$	[83], [123]
Composant	Pertes ou gain (dB)	Référence
$L_{fibre,inter-racks}$ (1 km)	0,2	[102]-[103]
$L_{connecteurs}$	1	[104]
$G_{OA(in-line)}$	20	[117]-[118]
L_{WSS}	5	[113], [115]

Tableau 4.2. Valeurs de sensibilité des récepteurs PIN et APD et pertes à considérer dans le bilan de liaison de l'architecture POPI+.

D'autre part, la valeur réelle de $P_{in,inter-racks}$ peut être calculée en considérant les pertes et les gains des composants présents dans la passerelle du rack (direction ascendante), selon l'équation suivante :

$$P_{in,inter-racks} = P_{Rx,POPI} - L_{BPF,inter-racks} + G_{OA(booster)} \quad (4.3.3)$$

Où $P_{Rx,POPI}$ est la puissance reçue requise pour assurer un BER de 10^{-12} pour le réseau POPI intra-rack (selon le Tableau 3.1). $L_{BPF,inter-racks}$ correspond aux pertes de 5 dB dues au BPF supplémentaire en bande L (placé au niveau de la passerelle du rack pour filtrer le bruit de l'OA). $G_{OA(booster)}$ est le gain de 20 dB de l'OA placé au niveau de la passerelle du rack pour amplifier la sortie de POPI+ intra-rack avant le coupleur $N + 1 : 1$ de la structure inter-racks.

L'équation 4.3.3 détermine la valeur de $P_{in,inter-racks}$, qui doit respecter la condition de l'équation 4.3.2 ($\check{P}_{in,inter-racks}$). Tout cas contraire indiquera que ce n'est pas possible d'établir la connexion souhaitée du fait des pertes de puissance trop élevées. Ainsi, les limites de $P_{in,inter-racks}$ correspondent à -4 et -9 dBm pour les récepteurs PIN et APD, respectivement. Enfin, nous présentons sur les figures 4.9(a) et (b) la puissance d'émission minimale du réseau intra-rack ($\check{P}_{Tx,POPI\ intra-racks}$) et la puissance minimale en entrée du réseau inter-racks ($\check{P}_{in,inter-racks}$) en fonction du nombre de serveurs par rack et de racks connectés, respectivement.

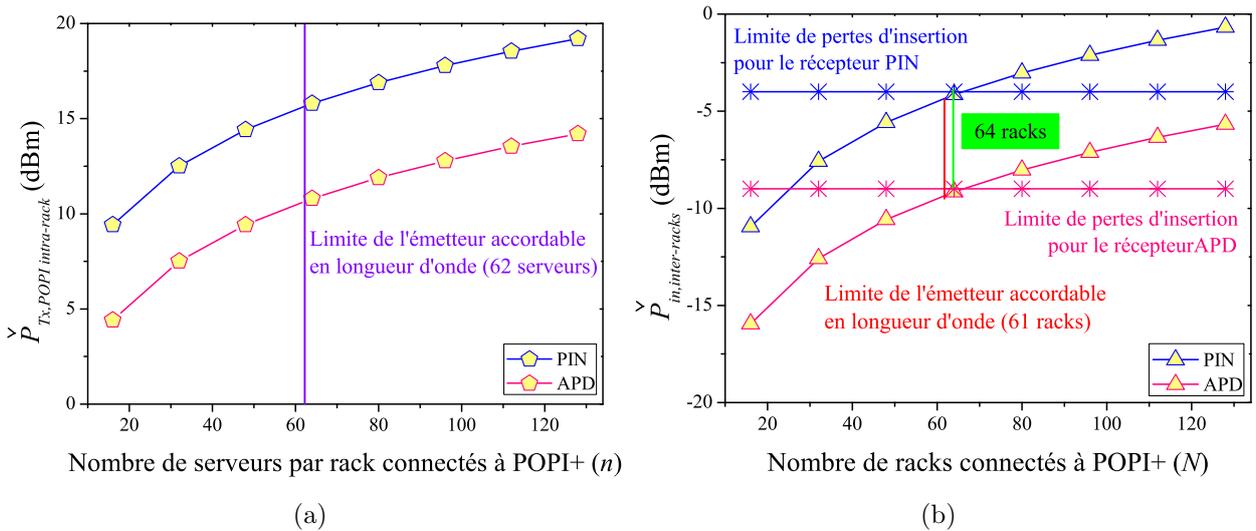


Figure 4.9. Capacité maximale du réseau POPI+ : (a) puissance d'émission minimale pour le réseau POPI intra-rack; (b) puissance minimale en entrée de la structure inter-racks.

Nous pouvons constater à partir de la Figure 4.9(a), que la puissance d'émission nécessaire pour maximiser le nombre de serveurs connectés par rack est de 10,8 et de 15,8 dBm pour les récepteurs PIN et APD, respectivement. Le nombre maximal de serveurs par rack est déterminé par la capacité du laser accordable en longueur d'onde en bande de transmission C, soit 64 canaux espacés de 50 GHz. Sachant que POPI+ réserve deux longueurs d'onde pour l'envoi et la réception de messages de *report* (λ_r) et *grant* (λ_g), le nombre maximal de serveurs par rack est de 62 pour des puissances d'émission supérieures ou égales à 10,8 ou

15,8 dBm selon le type de récepteur utilisé.

Nous constatons d'autre part, grâce à la Figure 4.9(b), que le nombre maximal de racks connectés à POPI+ en ne considérant que son bilan de liaison est de 64 racks. Cependant, ce chiffre est réduit à 61 dû au fait qu'il faut réserver trois longueurs d'onde de la bande de transmission L pour l'envoi du *report* (λ_r) et du *grant* (λ_g) ainsi que pour le *gateway* du DC (λ_G).

En conséquence, le réseau POPI+ permet de connecter jusqu'à 61 racks de 62 serveurs (soit 3782 serveurs) avec des connexions à 10 Gbit/s sous le format de modulation NRZ, en utilisant des récepteurs PIN ou APD.

4.4 Conclusion

Dans ce chapitre nous avons présenté le design d'une nouvelle architecture basée sur POPI et E-POPI sous le nom de POPI+, permettant d'augmenter le nombre de serveurs connectés. Cette nouvelle architecture considère deux étages de connexions basées sur TWIN. Les communications internes à chaque rack s'effectuent à l'aide d'un réseau POPI indépendant, où les longueurs d'onde peuvent être réutilisées. Les racks sont connectés grâce à une nouvelle structure, utilisant des WSS. Nous avons décrit le fonctionnement de cette nouvelle structure inter-racks et les éventuelles adaptations de POPI qui doivent être prises en compte sur les connexions intra-racks. Une variante de POPI+, sous le nom de Transfert POPI (T-POPI) a été également présentée, afin de privilégier les transferts inter-racks.

Nous avons ensuite étudié le bilan de liaison de POPI+ et nous avons établi le nombre maximal d'entités connectées lié aux pertes d'insertion des différents composants de cette architecture. Les résultats obtenus ont montré la possibilité de connecter jusqu'à 3782 machines pour un débit de 10 Gbit/s sous le format de modulation NRZ, en utilisant des récepteurs PIN ou APD. Il reste à évaluer cette capacité pour des débits supérieurs et avec des formats de modulation multiniveaux.

APPLICATION DU FORMAT PAM AUX RÉSEAUX OPTIQUES ÉTUDIÉS

5.1 Introduction

Ce chapitre a pour objet d'analyser les architectures d'interconnexion POPI, E-POPI et POPI+ du point de vue de la couche de transmission optique. Plus précisément, nous étudions le format de modulation d'amplitude PAM à M états afin de pouvoir augmenter le débit et l'efficacité spectrale de ces réseaux. Notre étude considère un volet de manipulation expérimentale, à des débits jusqu'à 18 Gbit/s, et un autre volet de simulation par ordinateur, avec des débits plus élevés (jusqu'à 112 Gbit/s).

La capacité maximale des réseaux POPI, E-POPI et POPI+ est estimée grâce aux résultats obtenus dans le deuxième volet combiné aux caractéristiques et aux limitations physiques des composants optiques présents sur chacun de ces réseaux.

5.2 Manipulations expérimentales

5.2.1 Banc de mesure

Nous présentons le schéma bloc du banc de mesure utilisé pour représenter un système PAM en détection directe sur la Figure 5.1. Il consiste en un émetteur composé par une onde porteuse provenant d'un analyseur de modulation optique Agilent Technologies' N4391A [124], une onde modulante provenant d'un générateur de signaux arbitraires Agilent Technologies' M8190A [125], suivi par un amplificateur électrique et un modulateur optique Mach-Zehnder. Nous utilisons 200 m de fibre optique monomode standard, un atténuateur variable et un filtre passe-bas optique. Les deux sorties d'un coupleur 10%/90% sont reliées respectivement à un analyseur de spectre optique et à une photodiode PIN qui est connectée à l'analyseur de modulation.

5.2.1.1 Emission

5.2.1.1.1 Onde porteuse

La porteuse optique est émise par l'analyseur de modulation, sa longueur d'onde et sa

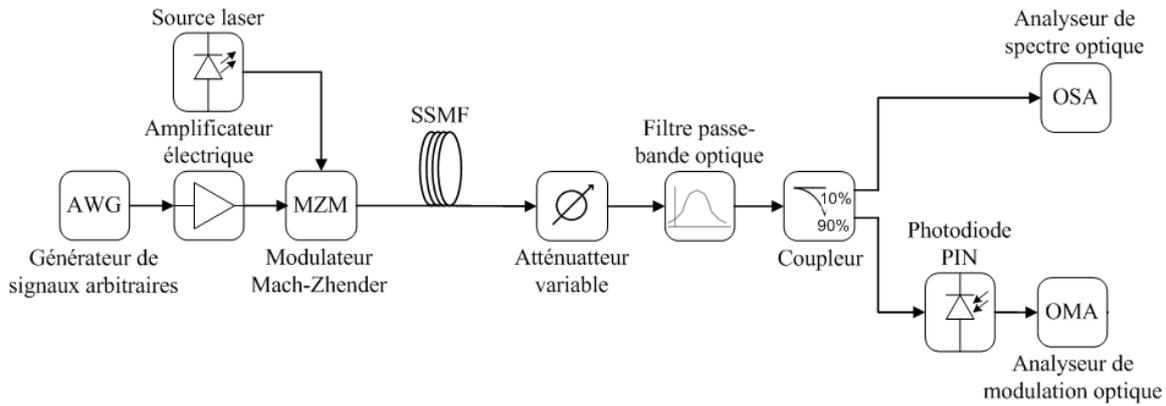


Figure 5.1. Schéma bloc du banc de mesure d'un système PAM en détection directe.

puissance sont fixées à 1550 nm et à 10 dBm, respectivement. Ce signal est représenté par une source laser sur le schéma de la Figure 5.1.

5.2.1.1.2 Onde modulante

L'onde modulante est générée grâce à l'AWG et au logiciel de simulation VPItransmissionMakerTM9.7. Le générateur de signaux arbitraires dont nous disposons peut générer plusieurs formats de modulation (NRZ, QAM, QPSK, DMT, etc.) avec IQtools de MathWorks®[126]. Cependant, cet outil ne nous permet pas de générer des signaux PAM. Mais, il est toujours possible d'importer ce signal sous forme de fichier .dat. Il suffit juste de respecter le taux d'échantillonnage du générateur de signaux arbitraires de 12 Géch/s. Le fichier .dat est obtenu en utilisant le logiciel de simulation optique VPItransmissionMakerTM9.7, en accord avec la Figure 5.2.

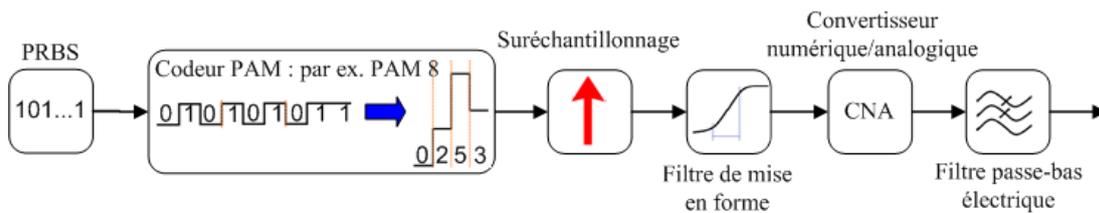


Figure 5.2. Génération du signal PAM sous VPItransmissionMakerTM9.7.

Le signal PAM généré sous VPItransmissionMakerTM9.7 est normalisé et transformé à posteriori en un signal bipolaire. Les amplitudes maximale, moyenne et minimale relatives correspondent donc à 1, 0 et -1, respectivement, selon la Figure 5.3. Le Tableau 5.1 résume les caractéristiques de chaque signal PAM en entrée du générateur de signaux arbitraires, concernant leur niveau de modulation, débit binaire, débit symbole et taux d'échantillonnage.

5.2.1.1.3 Modulateur optique

Nous disposons au laboratoire d'un modulateur Mach-Zehnder. Nous déterminons, avant

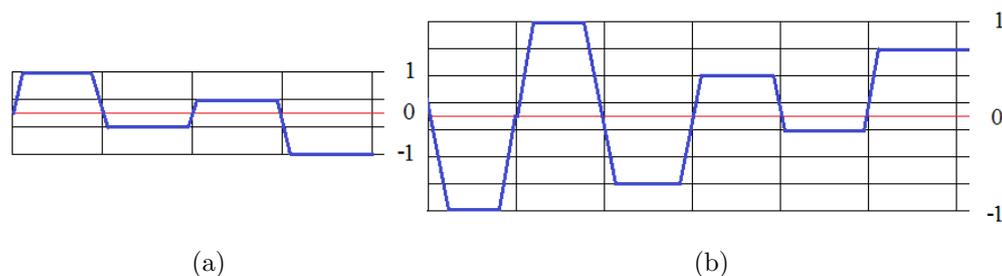


Figure 5.3. Signaux PAM 4 (a) et PAM 8 (b) bipolaires.

Niveau de modulation	Débit binaire (Gbit/s)	Débit symbole (GBaud)	Taux d'échantillonnage (Géch/symbole)
PAM 4	4	2	6
	6	3	4
	8	4	3
	12	6	2
PAM 8	6	2	6
	9	3	4
	12	4	3
	18	6	2

Tableau 5.1. Signaux PAM 4 et 8 en entrée du générateur de signaux arbitraires (générés sous VPItrans-missionMakerTM9.7).

de démarrer avec les mesures, le point de fonctionnement du modulateur, V_{MZM} , selon l'équation suivante :

$$V_{MZM} = \frac{V_{max} + V_{min}}{2} \quad (5.2.1)$$

où V_{max} et V_{min} représentent les tensions électriques appliquées au modulateur pour obtenir des puissances optiques maximale et minimale, respectivement.

Pour ce faire, le signal électrique provenant du générateur de signaux arbitraires (suivi par l'amplificateur électrique) est éteint et le signal optique provenant de la source laser est allumé et ajusté à 10 dBm. Ensuite, nous faisons varier la tension électrique appliquée au modulateur et nous récupérons la puissance optique affichée sur l'analyseur de spectre (résolution de 0,1 nm), nous présentons cette caractérisation sur la Figure 5.4.

A partir de ce graphique nous avons $V_{min} = 4,5$ V et $V_{max} = 11$ V, nous obtenons donc $V_{MZM} = 7,75$ V. A la sortie du générateur de signaux arbitraires le signal a une amplitude crête-à-crête de 0,7 V et le gain de l'amplificateur électrique est de 26 dB. Nous réduisons ce gain à 10 dB à l'aide de deux atténuateurs électriques de 10 et 6 dB, afin de pouvoir utiliser le modulateur dans une zone quasi linéaire. Avec ce gain le signal à l'entrée du modulateur a une amplitude crête-à-crête de 2,21 V ($0,7$ V $\times 10^{10/20}$). Par conséquent, les amplitudes

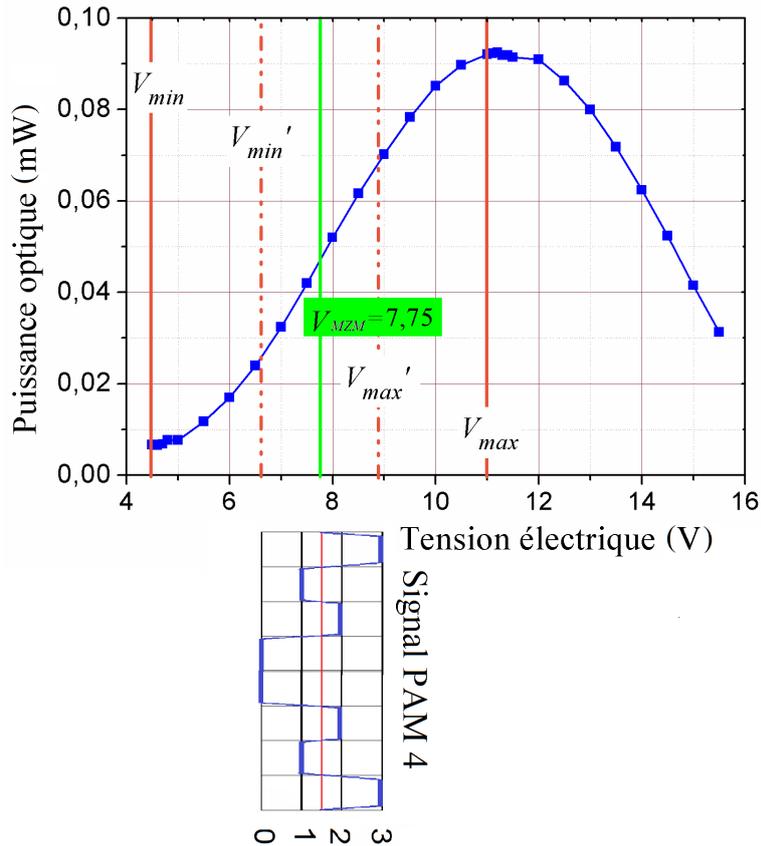


Figure 5.4. Caractérisation du MZM : puissance optique en fonction de la tension électrique appliquée.

maximale et minimale V_{max}' et V_{min}' du MZM sont respectivement de 8,86 et 6,64 V.

Ensuite, nous mesurons les pertes d'insertion du modulateur à l'aide de l'analyseur de spectre. Nous comparons les puissances optiques affichées sur l'analyseur de spectre de la chaîne optique présentée sur la Figure 5.1 (en absence de modulation) et du signal optique directement connecté à la fibre optique de transmission. Cette différence correspond aux pertes d'insertion de 7,4 dB.

5.2.1.2 Réception

5.2.1.2.1 Récepteur

Non seulement nous considérons la détection directe avec une photodiode PIN (de la Figure 5.1), mais encore nous incluons une pré-amplification avec un SOA (selon la Figure 5.5). La motivation de ceci étant l'amélioration de la sensibilité du récepteur, qui plus tard se traduira en une marge supplémentaire dans le bilan de liaison et nous permettra d'augmenter le nombre de machines connectées aux réseaux optiques. Nous utilisons un SOA avec un gain en zone linéaire de 18 dB et une puissance de saturation de 10 dBm.

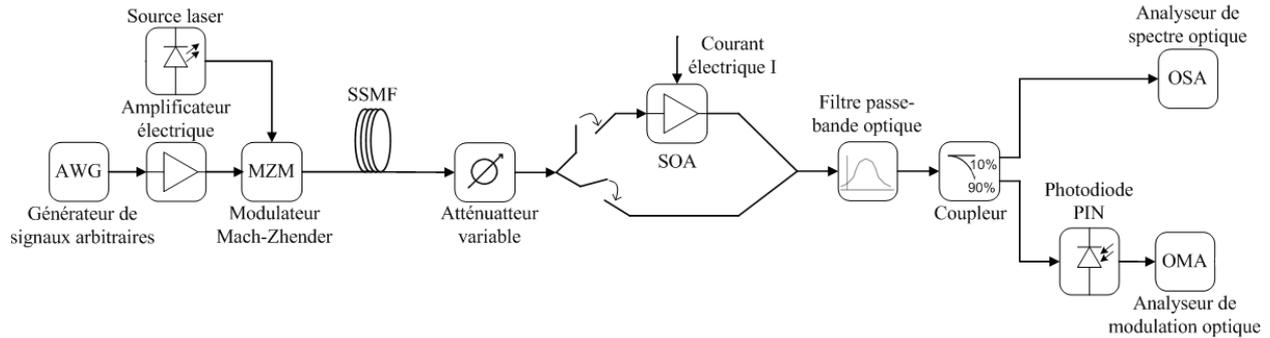


Figure 5.5. Schéma bloc du banc de mesure d'un système PAM en détection directe avec pré-amplification optique (SOA).

5.2.1.2.2 Filtrage

Nous disposons d'un filtre passe-bande optique reconfigurable Yenista OPTICS modèle XTM-50 [127]. Nous l'utilisons avec une largeur de bande à -3 dB de $1,5 \times$ le débit symbole du signal. Ce filtre présente des pertes d'insertion de 4 dB.

5.2.1.2.3 Analyseur de modulation optique

Nous analysons la qualité de transmission à l'aide de l'analyseur de modulation optique. Nous utilisons ce dernier en mode oscilloscope, pour voir les diagrammes de l'œil en électrique, les constellations et les calculs d'amplitude du vecteur d'erreur (EVM). Cet appareil, ne nous permettant pas l'analyse du format PAM, a été adapté en sélectionnant le format de modulation N-PSK (*N-Phase-Shift Keying*) sur l'application KEYSIGHT TECHNOLOGIES disponible sur cette machine. Nous pouvons choisir l'amplitude par rapport à l'origine et le déphasage entre les symboles ayant la même amplitude. La Figure 5.6 montre en coordonnées I (*in phase*) et Q (*quadrature*) les configurations PSK à N états choisies pour analyser les formats PAM 4 et 8.

Afin de détecter des symboles équidistants sur le signal reçu par la photodiode, nous avons choisi des amplitudes impaires consécutives à partir de « 1 » et un déphasage de 180° (ce qui nous permet d'avoir des symboles uniquement sur l'axe I).

5.2.1.2.4 EVM, SER et BER comme mesure de performance

Les résultats des manipulations expérimentales sont présentés en termes d'EVM. Nous comparons ces résultats avec des simulations par ordinateur à l'aide du logiciel VPItransmissionMakerTM9.7 en cosimulation avec le logiciel Matlab®R2014b. VPI nous permet d'avoir des résultats de simulation en termes de SER et Matlab des résultats en termes d'EVM. Dans cette section nous rappelons la définition de l'EVM et du SER et nous expliquons la façon dont nous nous servons de ces paramètres pour déterminer si la qualité du système optique est correcte ou pas.

Nous utiliserons la définition de l'EVM présentée dans [128]-[132] (Cf. Figure 5.7) :

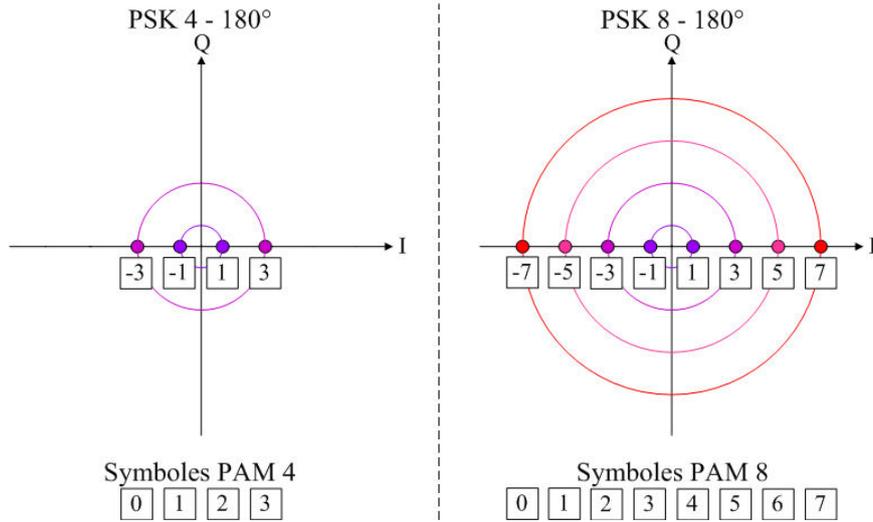


Figure 5.6. Configuration de l'application KEYSIGHT TECHNOLOGIES sur l'OMA pour détecter les formats PAM 4 et 8.

$$EVM = \frac{\sigma_{err}}{|E_{t,m}|}, \quad \sigma_{err}^2 = \frac{1}{I} \sum_{i=1}^I |E_{err,i}|^2, \quad E_{err,i} = E_{r,i} - E_{t,i} \quad (5.2.2)$$

Où σ_{err} représente l'écart type entre les symboles reçus $E_{r,i}$ et les symboles émis $E_{t,i}$. $|E_{t,m}|$ est l'amplitude maximale de la constellation de référence. I correspond au nombre total de symboles dans la séquence envoyée.

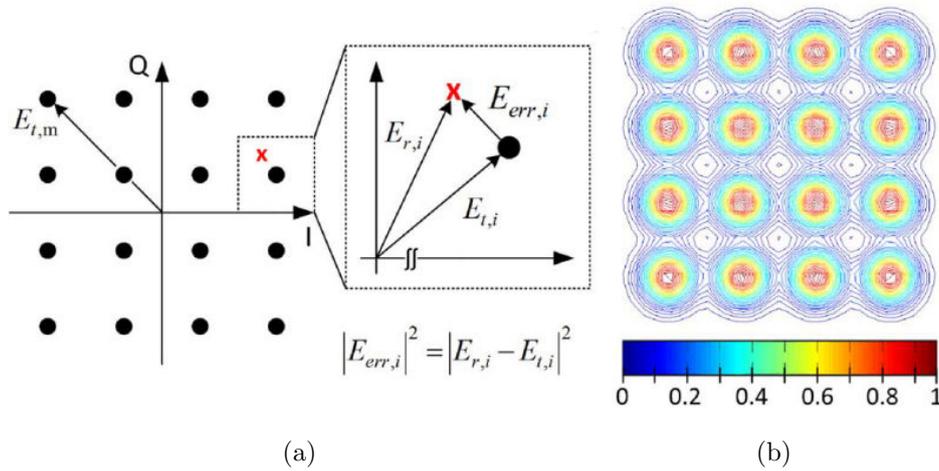


Figure 5.7. Constellation de QAM 16 pour le calcul de l'EVM : (a) constellation idéale ou de référence ; (b) constellation simulée avec du bruit gaussien (EVM=15%) [132].

En principe, dans le contexte des formats de modulation PAM à M états, le vecteur d'erreur E_{err} est un nombre réel puisque la partie imaginaire correspondant aux différences de déphasage est maintenant nulle. Cependant, expérimentalement, nous allons quand même

avoir du bruit électrique dans l'axe Q. Bruit que nous devons ajouter à l'équation 5.2.2 afin de pouvoir comparer les résultats expérimentaux aux simulations par ordinateur.

En outre, nous ne disposons pas au laboratoire du nécessaire pour synchroniser la séquence de symboles PAM envoyée avec celle captée par l'analyseur de modulation. Ce dernier donc a été configuré pour calculer l'EVM à l'aveugle. Autrement dit, au lieu de calculer l'EVM par rapport à des niveaux correspondant aux symboles envoyés, l'OMA déduit les niveaux de référence à partir de la distribution des tensions des différents échantillons reçus.

L'EVM calculé à l'aveugle sera donc inférieur à l'EVM de l'équation 5.2.2 puisque les niveaux de référence sont ajustés de façon à correspondre aux extrema locaux de la distribution des tensions. Nous montrons ceci sur la Figure 5.8. Nous prenons en compte cette modification dans l'étude par simulation par ordinateur.

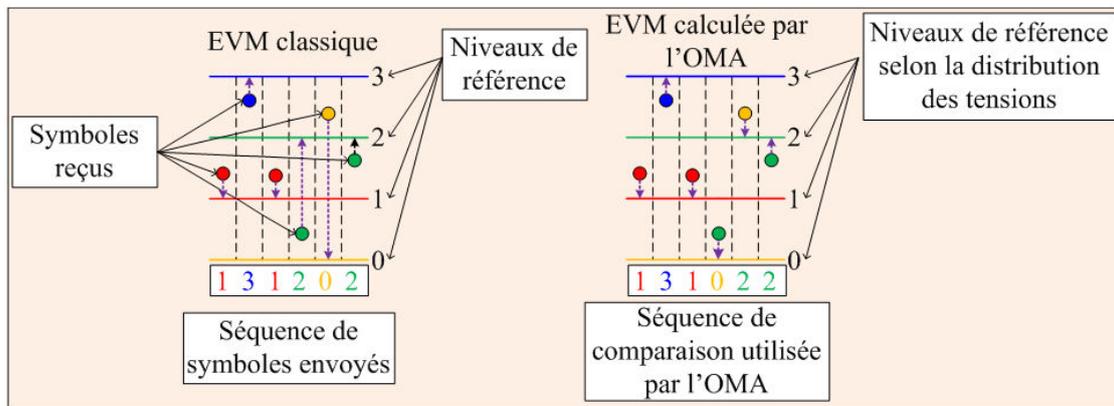


Figure 5.8. L'EVM classique contre l'EVM calculée par l'OMA.

Le taux d'erreur de symboles pour les formats de modulation PAM à M états est défini par l'équation suivante :

$$SER = \frac{1}{M} \sum_{j=0}^{M-1} P_{e,j} \quad (5.2.3)$$

Où $P_{e,j}$ représente la probabilité d'erreur du niveau j , qui peut être calculé en fonction du facteur Q (en prenant en compte uniquement la contribution des niveaux voisins au bruit) selon :

$$j = 0 : \quad P_{e,j} = \frac{1}{2} \operatorname{erfc} \left(\frac{Q_j}{\sqrt{2}} \right) \quad (5.2.4)$$

$$0 < j < M - 1 : \quad P_{e,j} = \frac{1}{2} \operatorname{erfc} \left(\frac{Q_{j-1}}{\sqrt{2}} \right) + \frac{1}{2} \operatorname{erfc} \left(\frac{Q_j}{\sqrt{2}} \right) \quad (5.2.5)$$

$$j = M - 1 : \quad P_{e,j} = \frac{1}{2} \operatorname{erfc} \left(\frac{Q_{j-1}}{\sqrt{2}} \right) \quad (5.2.6)$$

La fonction d'erreur complémentaire erfc étant définie par :

$$\operatorname{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-t^2} dt \quad (5.2.7)$$

et le facteur Q par :

$$Q_j = \frac{d_{j,j+1}}{\sigma_j + \sigma_{j+1}} \quad (5.2.8)$$

Avec $d_{j,j+1}$ égal à la distance entre la moyenne de symboles reçus deux niveaux consécutifs j et $j + 1$ et σ_j au bruit électrique du niveau j sur le niveau voisin $j + 1$, en accord avec la Figure 5.9.

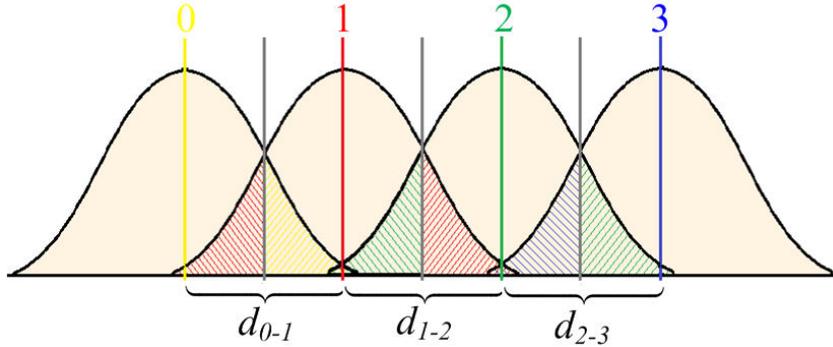


Figure 5.9. Fonctions de densité de probabilité de chaque niveau pour le format de modulation PAM 4.

$d_{j,j+1}$ peut être également défini par 5.2.9, si nous considérons que les moyennes des symboles reçus à chaque niveau PAM sont équidistantes et qu'en plus tous les symboles PAM ont la même probabilité d'être transmis.

$$d = d_{j,j+1} = \frac{\bar{P}_{max}R}{M-1} = \frac{\bar{P}_{moy}R}{2(M-1)} \quad (5.2.9)$$

Où \bar{P}_{max} représente la puissance moyenne des symboles « $M - 1$ » reçus ; \bar{P}_{moy} , la puissance moyenne de tous les symboles reçus et R , la détectivité du récepteur en A/W.

Nous pouvons ainsi réécrire 5.2.8 :

$$Q = Q_j = \frac{\bar{P}_{max}R}{2(M+1)\sigma_{th}} = \frac{\bar{P}_{moy}R}{4(M+1)\sigma_{th}} \quad (5.2.10)$$

Par conséquent, le SER correspond à :

$$SER = \frac{M-1}{M} \operatorname{erfc} \left(\frac{Q}{\sqrt{2}} \right) \quad (5.2.11)$$

Le taux d'erreurs binaires pour les formats de modulation PAM à M états peut être défini en fonction du SER par l'équation suivante :

$$\log_2(M) BER \geq SER \geq BER \quad (5.2.12)$$

Dans certains cas, notamment dans le cas où on utilise un codage Gray (où il n'existe qu'un seul bit différent entre chaque niveau consécutif) [133]-[134], le BER est égal au SER. Mais, en général, le SER représente une limite supérieure du BER. En conséquence, par la suite, nous présenterons nos résultats de simulation en termes de SER.

5.2.2 Résultats

La Figure 5.10 présente les résultats d'EVM obtenus en utilisant le schéma bloc de la Figure 5.1, en fonction de la puissance optique reçue, pour le format de modulation PAM 4, à une longueur d'onde fixe de 1550,12 nm, pour des débits symbole de 2, 3, 4 et 6 GBaud. Nous comparons sur cette figure les résultats d'EVM obtenus expérimentalement et par simulation par ordinateur. La simulation a été réalisée grâce au logiciel VPItransmissionMakerTM9.7 en co-simulation avec Matlab®R2014b. Dans la simulation nous avons utilisé les mêmes composants du schéma 5.1, avec les mêmes paramètres de simulation. Pour calculer l'EVM nous avons importé dans Matlab le signal reçu par la photodiode de VPI et programmé l'équation 5.2.2, qui prend comme séquence de comparaison la référence du niveau PAM 4 le plus proche au symbole reçu. Nous incluons sur la figure des valeurs de SER simulés, estimés par la méthode de Gauss [135].

D'une part, nous pouvons constater que, pour les quatre débits symbole étudiés, les valeurs d'EVM expérimentales et simulées sont très proches, surtout à des puissances optiques reçues supérieures à -16 dBm. Une erreur maximale de 5% peut être observée à une puissance optique reçue de -17 dBm, pour 6 GBaud. Ceci nous permet de valider notre outil de simulation pour la suite de ce manuscrit.

D'autre part, cette fois-ci, nous considérons un code correcteur d'erreurs binaires implanté en matériel (*Hard-decision FEC, HD-FEC*). Ce type de correcteur nous permet de tolérer des taux d'erreurs binaires avant correction de l'ordre de 10^{-3} [136]. Les codes correcteurs d'erreurs binaires implantés de façon logicielle (*Soft-decision FEC, SD-FEC*) peuvent tolérer des BER encore plus élevés, cependant ils introduisent des latences plus élevées [137] qui pourraient être contraignantes pour certaines applications internes au centre de données. Conformément l'équation 5.2.12 nous utiliserons la limite du HD-FEC pour déterminer le SER maximal admissible. Pour cette raison, nous extrayons depuis les courbes de SER, les valeurs de puissance optique reçue pour lesquelles le SER est égal à 10^{-3} et nous les utilisons pour obtenir les valeurs d'EVM expérimentale et par simulation (voir le Tableau 5.2).

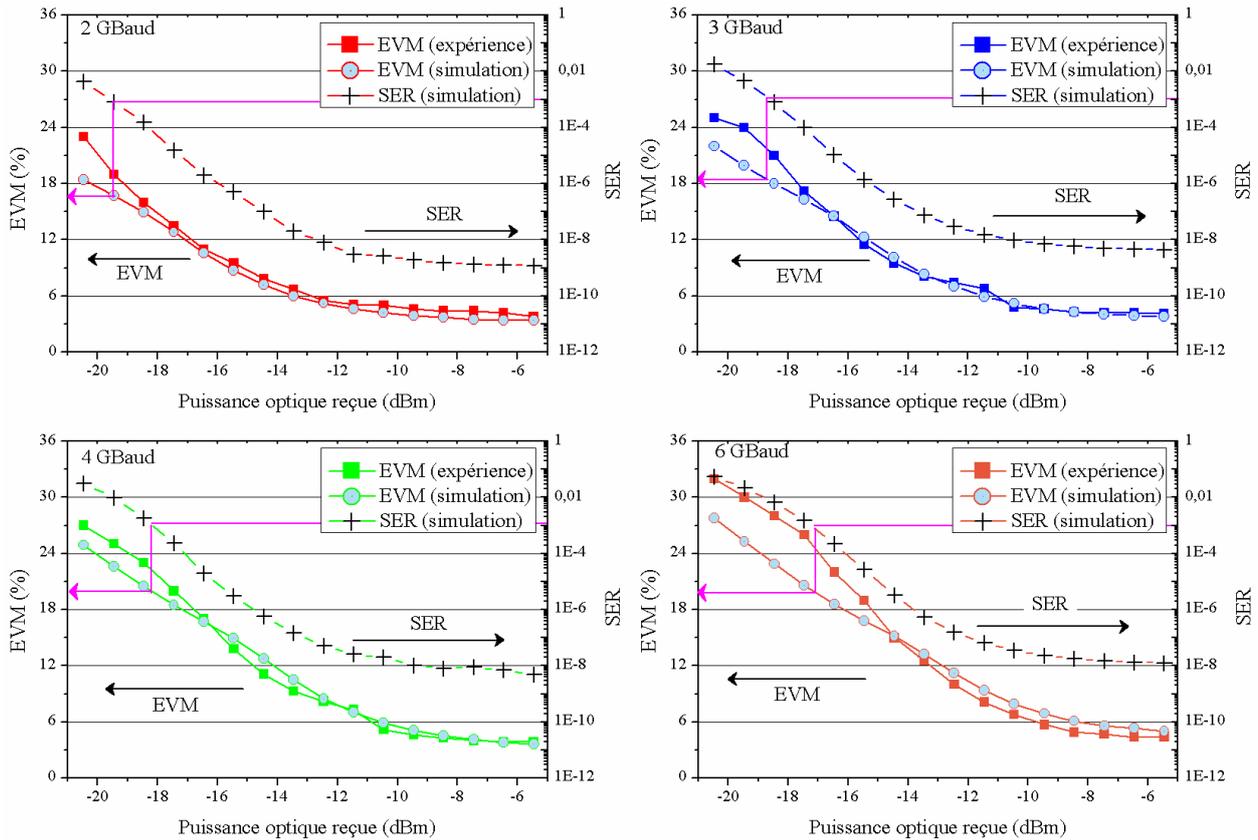


Figure 5.10. Résultats en termes d'EVM expérimental et par simulation, ainsi que de SER simulé, en fonction de la puissance optique reçue, pour le format de modulation PAM 4 à des débits symboles de 2, 3, 4 et 6 GBaud.

Paramètre \ Débit symbole (Gbaud)	2	3	4	6
Puissance optique nécessaire (dBm)	-19,4	-18,7	-18,2	-17
EVM expérimentale (%)	19,02	21,59	22,46	24,82
EVM par simulation (%)	16,65	18,37	19,87	19,87

Tableau 5.2. Puissance optique nécessaire pour un SER simulé de 10^{-3} , EVM expérimentale et par simulation à cette puissance, pour des débits symbole de 2, 3, 4, 6 GBaud.

Nous constatons grâce à ce tableau, que les valeurs d'EVM nécessaire pour un SER de 10^{-3} expérimentales ou par simulation ont tendance à augmenter avec le débit symbole.

La Figure 5.11 présente les résultats d'EVM obtenus en utilisant le schéma bloc de la Figure 5.1, en fonction de la puissance optique reçue, pour les formats de modulation PAM 4 et 8, à une longueur d'onde fixe de 1550,12 nm, pour des débits symbole de 2, 3, 4 et 6 GBaud.

Nous notons que l'ensemble des valeurs d'EVM diminuent considérablement lors du passage de PAM 4 à 8. Cet effet n'est pas surprenant, puisque le calcul à l'aveugle de l'EVM ne prend pas en compte des écarts supérieurs à la distance entre deux niveaux consécutifs.

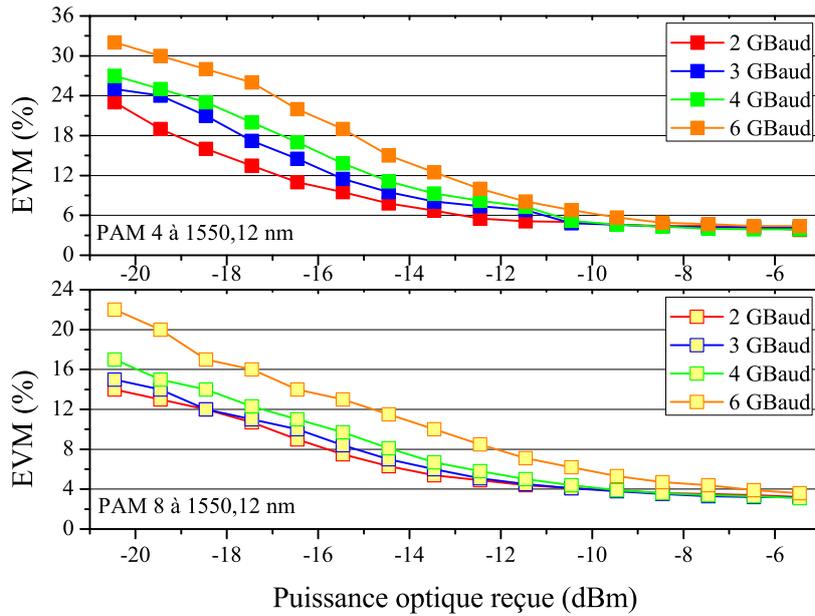


Figure 5.11. Résultats d'EVM expérimentale en fonction de la puissance optique reçue, pour les formats de modulation PAM 4 et 8, à une longueur d'onde fixe de 1550,12 nm, pour des débits symboles de 2, 3, 4 et 6 GBaud.

La Figure 5.12 présente les résultats d'EVM obtenus en utilisant le schéma bloc de la Figure 5.5, en fonction de la puissance optique reçue, pour les formats de modulation PAM 4 et 8, pour un débit symbole fixe de 6 GBaud, à des longueurs d'onde de 1550,12, 1520,25 et 1570,42 nm.

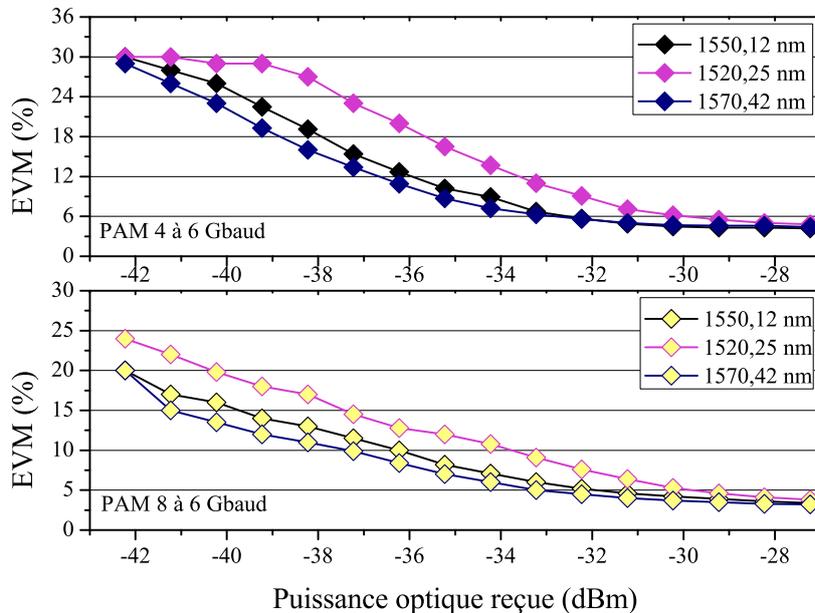


Figure 5.12. Résultats d'EVM expérimentale en fonction de la puissance optique reçue, pour les formats de modulation PAM 4 et 8, pour des longueurs d'onde de 1550,12, 1520,25 et 1570,42 nm, pour un débit symbole fixe de 6 GBaud.

Nous observons sur cette figure que le fait d'utiliser un SOA en réception va évidemment améliorer la sensibilité du récepteur. Nous mettons aussi en évidence l'impact de la dépendance du gain du SOA en fonction de la longueur d'onde.

Finalement, la Figure 5.13 présente les résultats obtenus toujours en utilisant le schéma bloc de la Figure 5.5, en fonction de la longueur d'onde, pour les formats de modulation PAM 4 et 8, pour un débit symbole fixe de 6 GBaud. Sur cette figure nous montrons la puissance optique affichée sur l'OSA et l'EVM sans et avec pré-amplification optique. Avec la pré-amplification optique nous ajustons l'atténuateur variable à 20 dB.

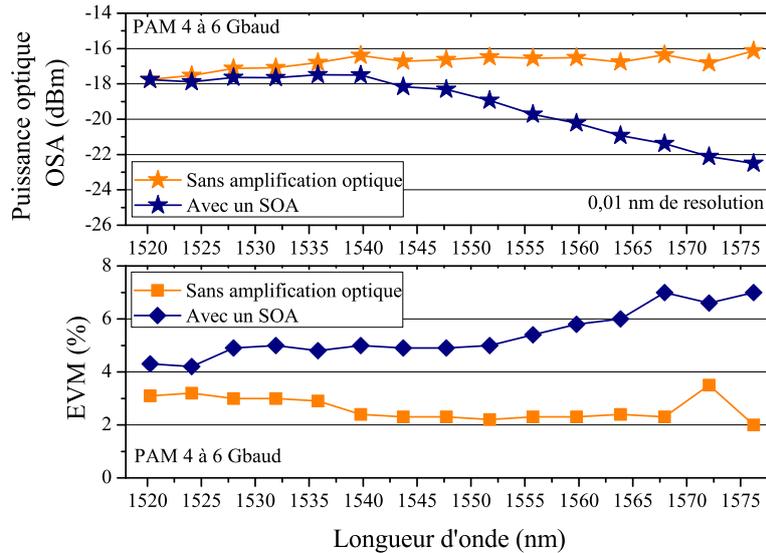


Figure 5.13. Résultats de puissance optique reçue au niveau de l'OSA et d'EVM expérimentale en fonction de la longueur d'onde, pour le format de modulation PAM 4 à 6 GBaud, sans et avec pré-amplification optique à l'aide d'un SOA.

La Figure 5.13 nous permet de voir à un même niveau de puissance l'effet de la variation du gain du SOA en fonction de la longueur d'onde.

5.3 Simulations sous VPItransmissionMaker 9.7

Cette section est dédiée à l'étude par simulation de liens optiques à différents débits (de 14 à 112 Gbit/s), sous les formats de modulation PAM 4 et 8, en utilisant trois types de récepteurs différents (PIN, APD et SOA-PIN). Les paramètres de simulations ont été choisis de façon plutôt pessimiste afin de simuler les pires conditions possibles.

5.3.1 Schémas de simulation

Le schéma de simulation de base est présenté sur la Figure 5.14. La transmission optique de 5 canaux WDM adjacents est mise en place afin de considérer les effets de diaphonie. Les résultats que nous présenterons par la suite correspondent au canal central (canal 3). Les

résultats obtenus en ayant utilisé ce schéma nous permettrons éventuellement d'estimer le dimensionnement des réseaux POPI, E-POPI et POPI+ intra-rack.

Nous modifions le schéma de simulation de la Figure 5.14. Le schéma de modulation modifié est présenté sur les figures 5.15 (a) et (b). Sur ces figures nous avons inclus un ou deux amplificateur(s) optique(s) EDFA. Ces schémas de simulation vont être utilisés pour le dimensionnement de POPI+ inter-racks.

Nous utilisons, dans un premier temps, le schéma de simulation de la Figure 5.15(a), avec un EDFA qui représente l'amplificateur optique de ligne du réseau POPI+ inter-racks placé entre le coupleur passif et la cascade de WSS (voir Figure 4.2).

Dans un deuxième temps, le schéma de simulation de la Figure 5.15(b) (chaîne complète de POPI+) est considéré, avec un deuxième EDFA qui va représenter l'amplificateur optique en mode *booster* au niveau du *gateway* du réseau POPI+ intra-rack (voir Figure 4.3).

Afin de pouvoir régler la puissance totale à l'entrée du premier EDFA (figures 5.15 (a) et (b)), nous utilisons le module *AmpSysOpt*, qui correspond à un amplificateur optique idéal (sans bruit ASE), suivi par un atténuateur. De même, nous utilisons un deuxième atténuateur avant le deuxième EDFA (Figure 5.15(b)).

La puissance en sortie de l'amplificateur optique idéal a été fixée à 5 mW et les deux atténuateurs sont réglés à 20 dB. Les deux EDFA possèdent les mêmes caractéristiques : la longueur de la fibre dopée à l'erbium est de 13 m, la durée de vie des ions d'erbium dans le niveau excité est de 10 ms, la pompe en contre-propagation à 980 nm a une puissance de 30 mW. Ces paramètres donnent un gain en zone linéaire de 23,8 dB et une puissance de saturation de 6,8 dBm (-0,2 dBm par canal WDM). La puissance totale à l'entrée de chaque EDFA est de -13 dBm, puissance pour laquelle le gain est de 20,3 dB.

5.3.1.1 Emission

5.3.1.1.1 Onde porteuse

Le module *LaserCW_DSM* est ajusté en longueur d'onde et puissance à 1550,12 nm (pour le canal central, les canaux sont espacés de 0,8 ou 0,4 nm selon le débit et le format de modulation) et à 10 dBm, respectivement. Les paramètres de largeur de bande et de bruit relatif d'intensité (*RIN : Relative Intensity Noise*) ont été fixés à 5 MHz et à -160 dB/Hz avec une puissance moyenne de 10 dBm, respectivement.

5.3.1.1.2 Onde modulante

L'onde modulante est générée grâce à la galaxie (ensemble de modules VPI indépendants, avec une fonctionnalité déterminée) *Source PAM*. Chaque canal a une source PAM comme celle qui a été décrite dans la section 5.2.1 (Figure 5.2). Cette fois-ci la longueur de la séquence est de $2^{10} - 1$ bits.

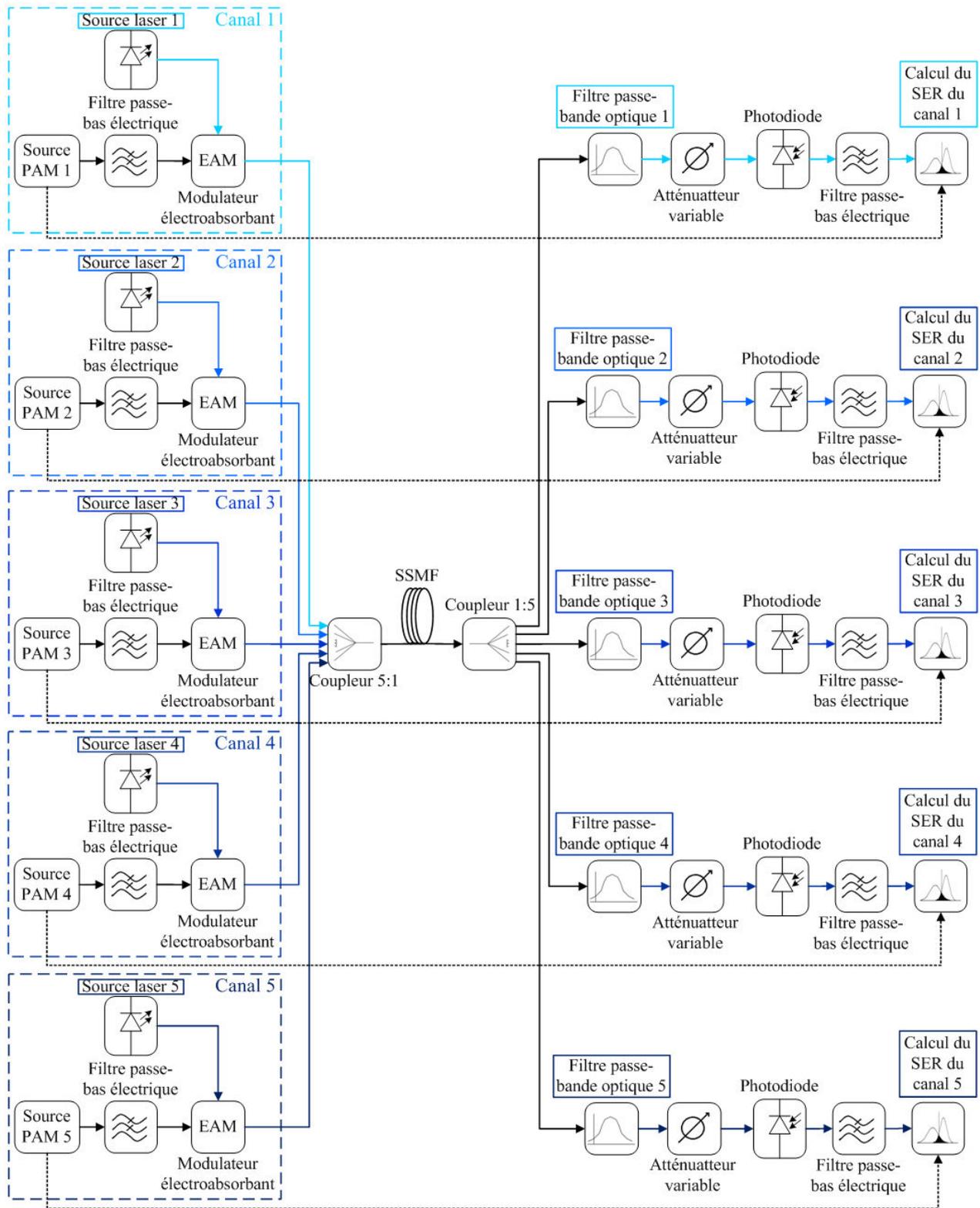


Figure 5.14. Schéma de simulation des réseaux POPI, E-POPI et POPI+ intra-rack : transmission optique de 5 canaux WDM avec le format de modulation PAM (4 ou 8).

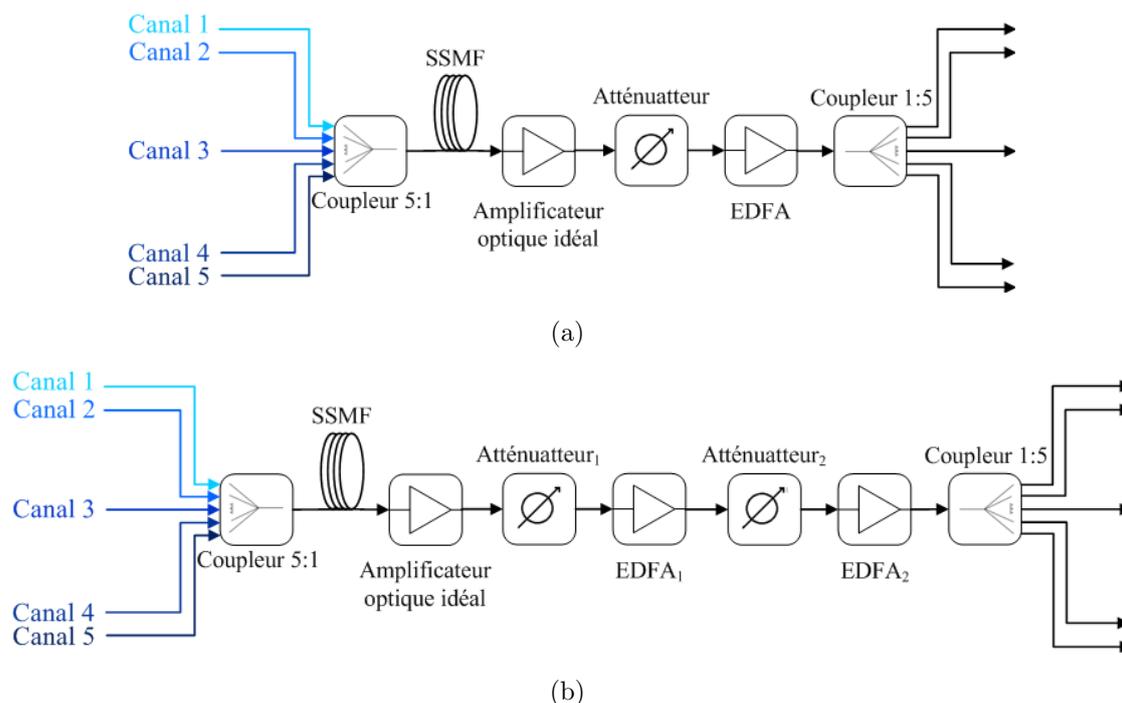


Figure 5.15. Schéma de simulation modifié : insertion d'un ou deux EDFA sur la ligne de transmission optique pour (a) et (b), respectivement. Ces schémas sont utilisés pour simuler le réseau POPI+ inter-racks.

5.3.1.1.3 Modulateur optique

Nous utilisons un modulateur électro-absorbant. Ce choix est justifié principalement par les raisons suivantes :

- Possibilité d'intégration avec les émetteurs, ce qui réduit considérablement les coûts du réseau [31]-[32], [138]-[139].
- Efficacité énergétique supérieure à celle des modulateurs Mach-Zehnder [139].

Nous utilisons le module *ModulatorEA_Measured*, dont la fonction de transfert est présentée sur la Figure 5.16. Cette fonction de transfert, ayant des zones linéaires très réduites au centre de la courbe, nous limite en termes d'amplitude maximale de la tension électrique à appliquer. Ceci est contraignant, car le signal multiniveaux est plus sensible au bruit et aux interférences inter-symboles (pour une même taux d'extinction). C'est pour cette raison que nous proposons par la suite une pré-distorsion du modulateur, qui nous permettra de travailler avec une tension électrique plus élevée, tout en ayant une réponse globale à-peu-près linéaire avec un taux d'extinction acceptable (d'environ 13 dB).

Les paramètres *Bias* et *Drive* vont définir les amplitudes maximale et minimale du signal électrique PAM, selon : $V_{max}(V) = Drive + Bias$ et $V_{min}(V) = Bias$, respectivement. Le point de fonctionnement de l'EAM est donc donné par l'équation suivante :

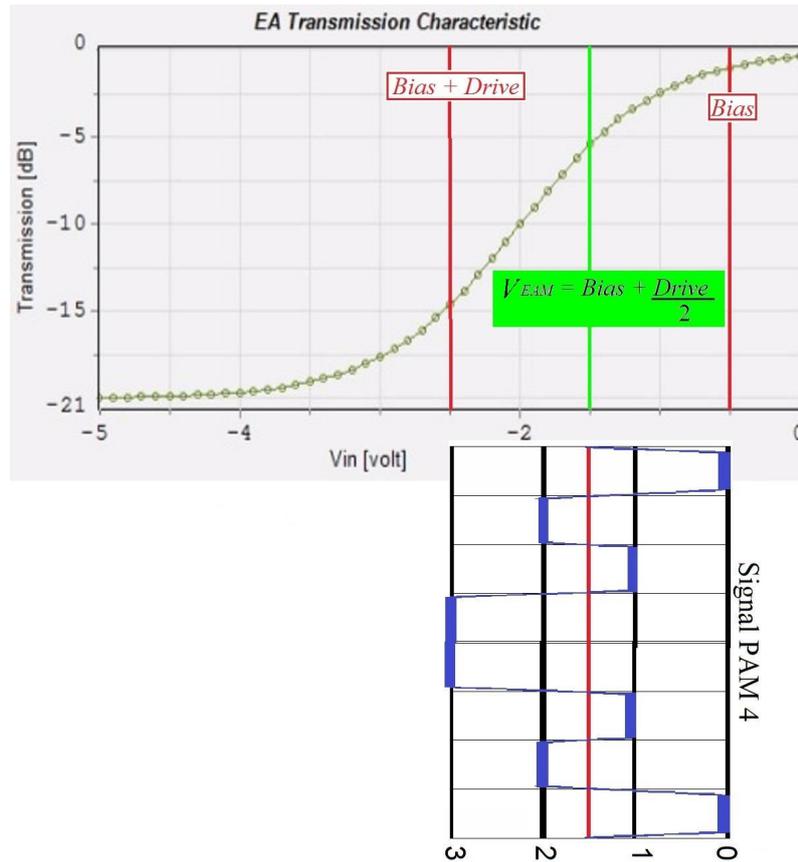


Figure 5.16. Fonction de transfert du module *ModulatorEA_Measured* de VPI : puissance optique en fonction de la tension électrique appliquée (figure de VPItransmissionMakerTM9.7 modifiée afin de visualiser les paramètres *Bias* et *Drive*, ainsi que le signal PAM 4 en entrée du modulateur).

$$V_{EAM} = V_{min} + \frac{V_{max} - V_{min}}{2} = Bias + \frac{Drive}{2} \quad (5.3.1)$$

5.3.1.1.4 Filtrage

Le signal provenant de la source PAM est filtré grâce à un filtre passe-bas électrique, ce dernier a une fonction de transfert de type Bessel d'ordre 4 et une largeur de bande à -3 dB de $0,75 \times$ débit symbole.

5.3.1.2 Réception

5.3.1.2.1 Récepteur

Encore une fois nous considérons la détection directe avec une photodiode PIN, sans ou avec une pré-amplification grâce à un SOA. Nous incluons cette fois-ci des résultats avec une photodiode à avalanche. La Figure 5.17 présente les différents types de récepteur. Le module *Photodiode* de VPI peut être configuré en mode « PIN » ou « APD », nous l'avons donc utilisé pour simuler ces trois cas. Pour le troisième cas, nous avons inséré le module *AmpBlackBox-*

SOA avant la photodiode PIN ; ce module permet de simuler la pré-amplification du SOA en utilisant des caractéristiques de gain et de facteur de bruit réelles. Le Tableau 5.3 montre les paramètres de chaque récepteur. Les résultats sont exprimés en termes de SER estimé par la méthode de Gauss [135].

Paramètre	Valeur numérique
Délectivité ^a	0,9/0,9/0,8/0,65 A/W ^d
Courant d'obscurité ^a	10 nA
Facteur multiplicatif d'avalanche ^b	10
Coefficient d'ionisation ^b	0,45
Bruit thermique ^a	20 pA/Hz ^{1/2}
Bruit de grenaille ^a	« ON » ^e
Gain du SOA ^c	20 dB
Puissance de saturation du SOA ^c	10 dBm

^aParamètres applicables aux récepteurs PIN, APD et SOA-PIN.

^bParamètres applicables seulement au récepteur APD.

^cParamètres applicables seulement au récepteur SOA-PIN.

^dValeurs de délectivité pour des débits symboles de 7, 14, 28 et 56, respectivement.

^eLe module *Photodiode* permet de configurer le bruit de grenaille en « ON » ou « OFF »

Tableau 5.3. Paramètres considérés pour les récepteurs PIN, APD et SOA-PIN.

5.3.1.2.2 Filtrage

Deux ou trois filtres en réception sont requis :

- Un filtre passe-bas électrique pour filtrer le signal provenant de la photodiode (PIN ou APD), avec les mêmes caractéristiques que le filtre passe-bas électrique en émission.
- Un filtre passe-bande optique, pour filtrer les canaux WDM, nous considérons des filtres trapézoïdaux avec une largeur de bande minimale de $1,5 \times$ débit symbole et une atténuation de 20 dB (avec une pente maximale de 800 dB/nm [56]).
- Enfin dans le cas du récepteur SOA-PIN, un troisième filtre passe-bande optique (avec les même caractéristiques que le filtre précédent) permet de filtrer le bruit d'émission spontanée amplifiée du SOA.

5.3.2 Pré-compensation de la distorsion

La fonction de transfert non-linéaire de l'EAM suggère qu'un signal PAM ayant des niveaux équidistants à l'entrée du modulateur vont générer, en sortie du modulateur, un signal modulé ayant des niveaux PAM distordus. Le diagramme de l'œil en sortie de l'EAM aura des yeux rétrécis ou élargis aux bords ou au centre de la figure. La méthode de pré-distorsion que nous décrivons ci-dessous a pour but de corriger cet effet.

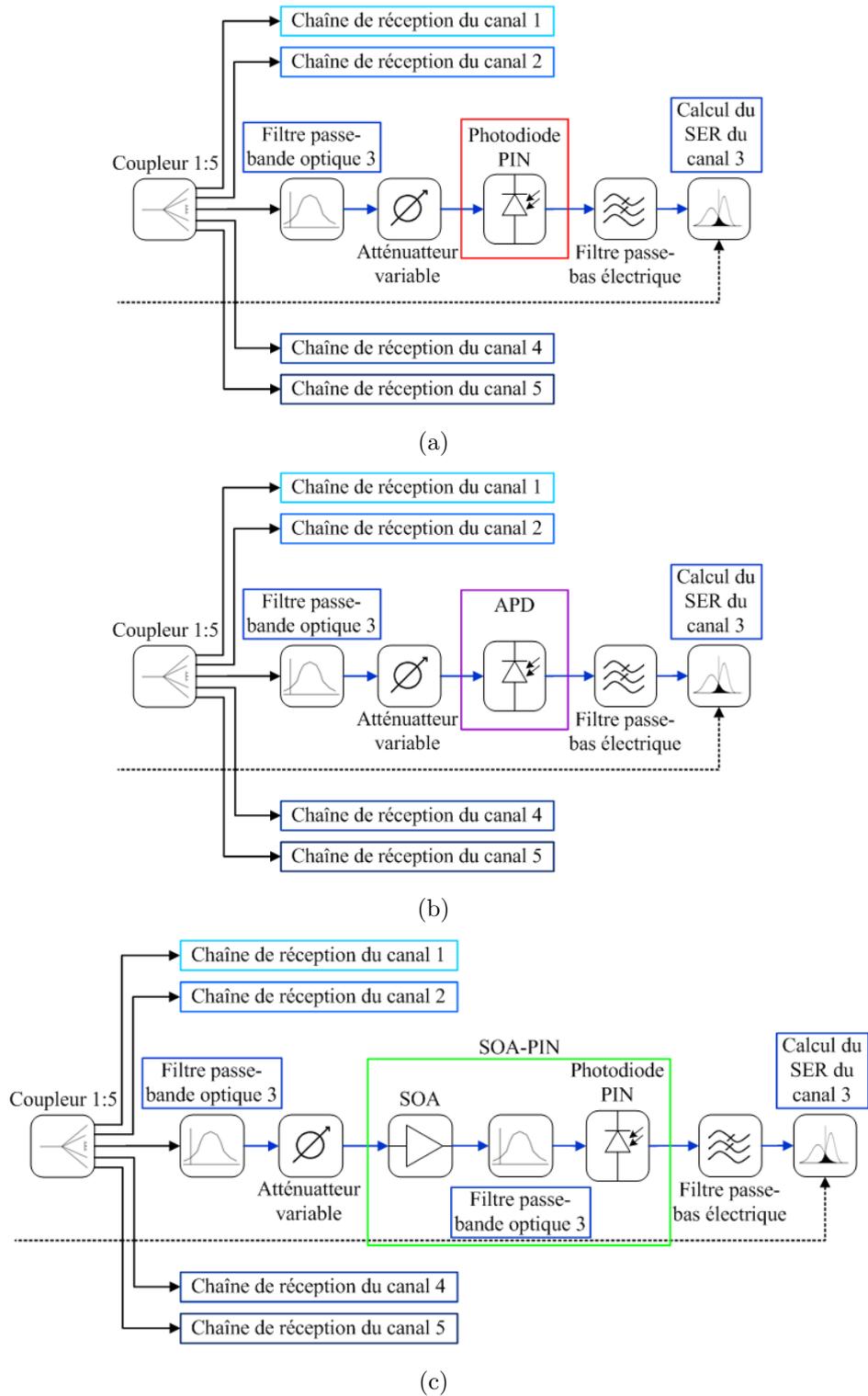


Figure 5.17. Schéma de simulation avec différents types de récepteurs optique : (a) photodiode PIN; (b) photodiode APD; (c) SOA + photodiode PIN.

- 1. Choisir une région délimitée par V_{max} et V_{min} afin de borner les non-linéarités de la fonction de transfert du modulateur, tout en réduisant le taux d'erreurs de symboles

(SER). Ceci doit être fait en faisant varier les paramètres de *Bias* et de *Drive*. En réduisant cette région nous limitons les non-linéarités en dépit de la distance entre les symboles. Autrement dit, le signal PAM devient alors plus vulnérable au bruit.

- 2. Redéfinir les niveaux PAM du signal électrique à l'entrée du modulateur en additionnant ou soustrayant une constante Δ à certains des niveaux PAM équidistants, selon le Tableau 5.4. Δ est choisie en minimisant le SER.
- 3. Recommencer les deux étapes précédentes jusqu'à trouver un minimum du SER.

Niveaux équidistants de PAM 4							
0,01 ^a		1/3		2/3			1
Niveaux pré-distordus de PAM 4							
0,01 ^a		1/3		2/3- Δ			1
Niveaux équidistants de PAM 8							
0,01 ^a	1/7	2/7	3/7	4/7	5/7	6/7	1
Niveaux pré-distordus de PAM 8							
0,01 ^a	1/7+ Δ	2/7+ Δ	3/7	4/7- Δ	5/7- Δ	6/7- Δ	1

^aLe premier niveau (« 0 ») a pour valeur numérique 0,01, afin d'assurer un taux d'extinction de 13 dB en sortie du modulateur.

Tableau 5.4. Niveaux équidistants et prédistordus des signaux PAM 4 et 8.

La Figure 5.18 montre le diagramme de l'œil à différents points du schéma de simulation. La 1^{ère} ligne (figures 5.18(a), (b), (c) et (d)) présente des résultats pour PAM 4 ; la 2^{ème} ligne (figures 5.18(e), (f), (g) et (h)), pour PAM 8. Sur ces figures les niveaux PAM sont inversés (par rapport aux niveaux présentés sur la Figure 5.4), pour être en accord avec la fonction de transfert de l'EAM. Les figures 5.18(a) et (e) montrent le signal électrique PAM à l'entrée du modulateur avec des niveaux PAM équidistants (sans avoir appliqué aucune méthode de pré-distorsion). Les figures 5.18(b) et (f) montrent l'effet de distorsion introduit par l'EAM, le signal optique PAM à la sortie du modulateur présente des niveaux PAM distordus. Les figures 5.18(c) et (g) montrent le signal électrique PAM à l'entrée du modulateur après être passé par l'étage de pré-distorsion. Finalement, les figures 5.18(d) et (h) montrent le signal optique PAM à la sortie de l'EAM, en ayant utilisé les signaux des figures 5.18(c) et (g) en entrée.

Comme prévu, les diagrammes de l'œil après passage par l'EAM présentent des yeux rétrécis pour les niveaux bas et hauts et des yeux élargis pour les niveaux moyens. En outre, les diagrammes de l'œil en entrée de l'EAM, ayant subi la méthode de pré-distorsion, présentent l'allure inverse : des yeux élargis pour les niveaux bas et hauts et des yeux rétrécis pour les niveaux moyens.

La compensation de la distorsion introduite par l'EAM est claire pour PAM 4, mais un peu moins évidente pour PAM 8. Cela est normal, car nous avons introduit une variation Δ unique

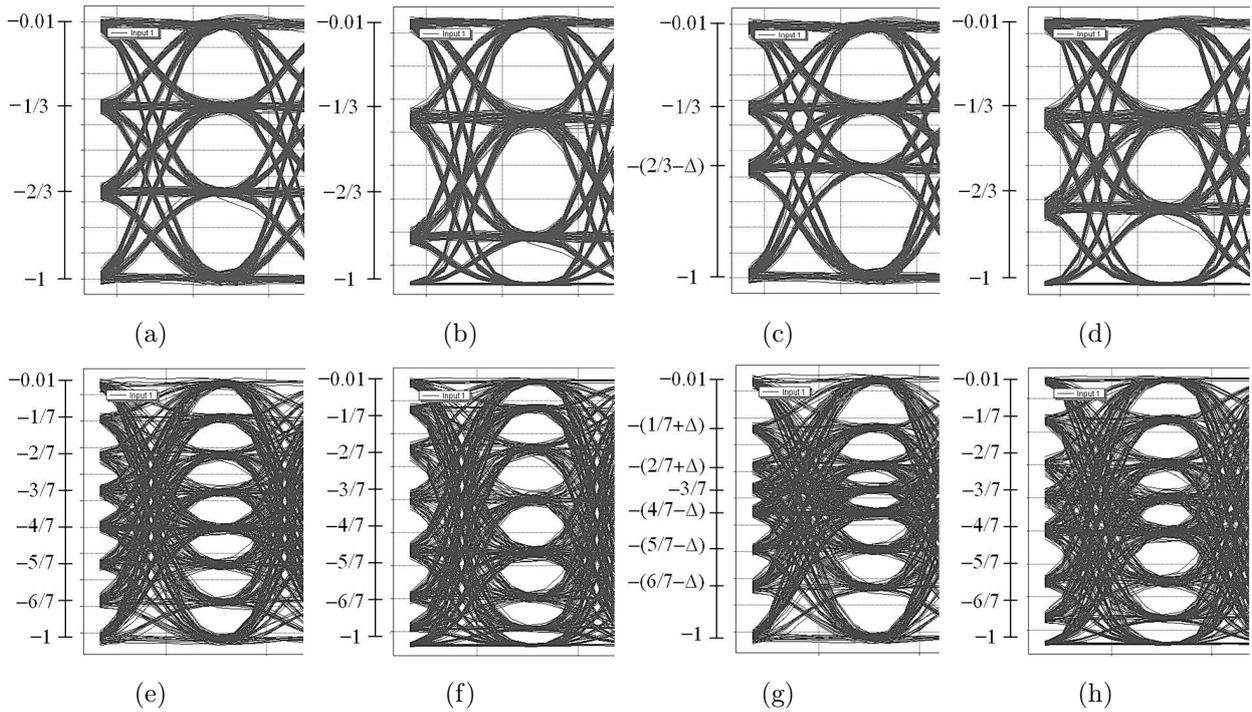


Figure 5.18. Diagramme de l'œil de PAM. Les lignes représentent le niveau de modulation M : PAM 4 ((a), (b), (c) et (d)) et PAM 8 ((e), (f), (g) et (h)). Les figures sur les colonnes impaires ((a), (c), (e) et (g)) et paires ((b), (d), (f) et (h)) représentent le signal PAM électrique et le signal PAM optique à l'entrée et à la sortie du modulateur optique, respectivement. Les figures sur les 1^{ère} et 2^{ème} colonnes ((a), (b), (e) et (f)) représentent des signaux PAM sans méthode de pré-distorsion ; sur les 3^{ème} et 4^{ème} colonnes ((c), (d), (g) et (h)), avec la méthode de pré-distorsion décrite ci-dessus. Paramètres utilisés : $Bias = -0,4/-0,3$ V, $Drive = -2/-1,95$ V et $\Delta = 0,095/0,05$ pour PAM 4/8.

pour certains niveaux PAM. En toute rigueur, nous pourrions améliorer la performance de notre méthode en utilisant des variations ($\Delta_0, \Delta_1, \dots, \Delta_{3ou7}$) propres à chaque niveau. Ceci ajouterait un nouveau degré de liberté par niveau, arrivant à des résultats de compensation beaucoup plus précis. Nonobstant, le temps de calcul augmenterait considérablement.

Par exemple, avec la méthode originale, pour trouver les valeurs optimales de $Bias$, $Drive$ et Δ nous considérons des plages de $[-1; 0]$, $[-2,5; -1,5]$ et $[0; 0,015]$ aux pas de 0,05, 0,05 et 0,0005, respectivement. Nous faisons un premier balayage en deux dimensions pour $Bias$ et $Drive$ (1. de la méthode), c'est-à-dire, 20×20 simulations d'environ 2 minutes de temps de calcul chacune. Un deuxième balayage en une dimension pour Δ (2. de la méthode) est prévu, l'équivalent de 30 simulations de 2 minutes. Ensuite, deux balayages de deux et une dimensions vont être enchainés, en limitant les plages à la moitié, et ainsi de suite (3. de la méthode). Le temps de simulation de du premier balayage est d'environ 860 minutes, soit un temps maximal d'au plus 1680 minutes (soit 28 heures). Par contre, en pratique, un SER minimum a été trouvé au bout de deux itérations, soit 21 heures et 30 minutes. Nous n'avons pas proposé de balayages en 3 dimensions, car le temps de simulation aurait été beaucoup

plus élevé : $20 \times 20 \times 30 = 12000$ minutes (soit 8 jours et 8 heures).

Ainsi, le fait de considérer des variations Δ propres à chaque niveau, aurait signifié un balayage de 4 ou 8 dimensions (au lieu d'une seule) pour le deuxième pas de la méthode. Ceci aurait augmenté le temps de simulation de ce balayage de 30 à 30^4 ou $30^8 \times 2$ minutes, soit de 60 à $1,6 \times 10^6$ ou $1,2 \times 10^{12}$ minutes. Par conséquent, nous nous contentons des résultats obtenus grâce à la méthode proposée.

Cependant, cette méthode n'a été conçue que pour compenser la distorsion provenant d'une seule source, l'EAM. C'est-à-dire, elle est applicable pour les récepteurs PIN ou APD et non pas pour le récepteur SOA-PIN. Ce dernier constitue une deuxième source de distorsion causée par la saturation du SOA.

Afin de visualiser cet effet, la Figure 5.19 présente les diagrammes de l'œil à différents points du schéma de simulation en utilisant le récepteur SOA-PIN. La 1^{ère} ligne (figures 5.19 (a), (b) et (c)) présente des résultats pour PAM 8 avec une puissance optique reçue de -15 dBm ; la 2^{ème} ligne (figures 5.19 (d), (e), et (f)), de -10 dBm. Les figures 5.19 (a) et (d) montrent le signal électrique PAM à l'entrée du modulateur ayant passé par la méthode de pré-distorsion proposée précédemment. Les figures 5.19 (b) et (e) montrent le signal optique PAM à la sortie de l'EAM, en ayant utilisé (a) et (d) en entrée. Les figures 5.19 (c) et (f) montrent le signal électrique PAM à la sortie du récepteur SOA-PIN.

Le diagramme de l'œil après passage par le SOA-PIN est bien distordu pour les deux puissances reçues étudiées. Cette distorsion n'a pas du tout la même allure que celle introduite par le modulateur optique. Le SOA au lieu de rétrécir ou élargir les yeux des niveaux extérieurs ou moyens, va rétrécir les yeux des niveaux bas (en haut de chaque figure) et élargir les yeux des niveaux hauts (en bas de chaque figure). Cet effet est accentué en baissant la puissance reçue de -10 à -15 dBm. Il est donc indispensable de compenser cette deuxième source de distorsion.

D'une part, la méthode de compensation de l'EAM expliquée précédemment n'est plus totalement efficace. La distorsion introduite par le SOA n'étant pas symétrique, pour compenser la distorsion provenant de l'ensemble EAM + SOA-PIN il faut reconsidérer l'utilisation de variations Δ différentes pour chaque niveau. Pour le format de modulation PAM 4, la méthode originale arrive à compenser la distorsion introduite par l'ensemble EAM + SOA-PIN de manière à-peu-près satisfaisante (le SOA-PIN remplace la photodiode PIN ou APD et le SER est minimisé), car la valeur Δ n'intervient que sur un seul niveau. Alors que pour le format de modulation PAM 8, la méthode de compensation rajoute encore plus de distorsion que l'ensemble EAM + SOA-PIN (le SER est encore plus élevé). L'adaptation des variations Δ propres à chaque niveau devient donc nécessaire. Mais, d'autre part, la contrainte du temps de simulation pour faire des balayages à plusieurs dimensions, empêche d'utiliser des variations Δ différentes pour chaque niveau.

En conséquence, nous proposons une solution hybride qui considère une valeur unique

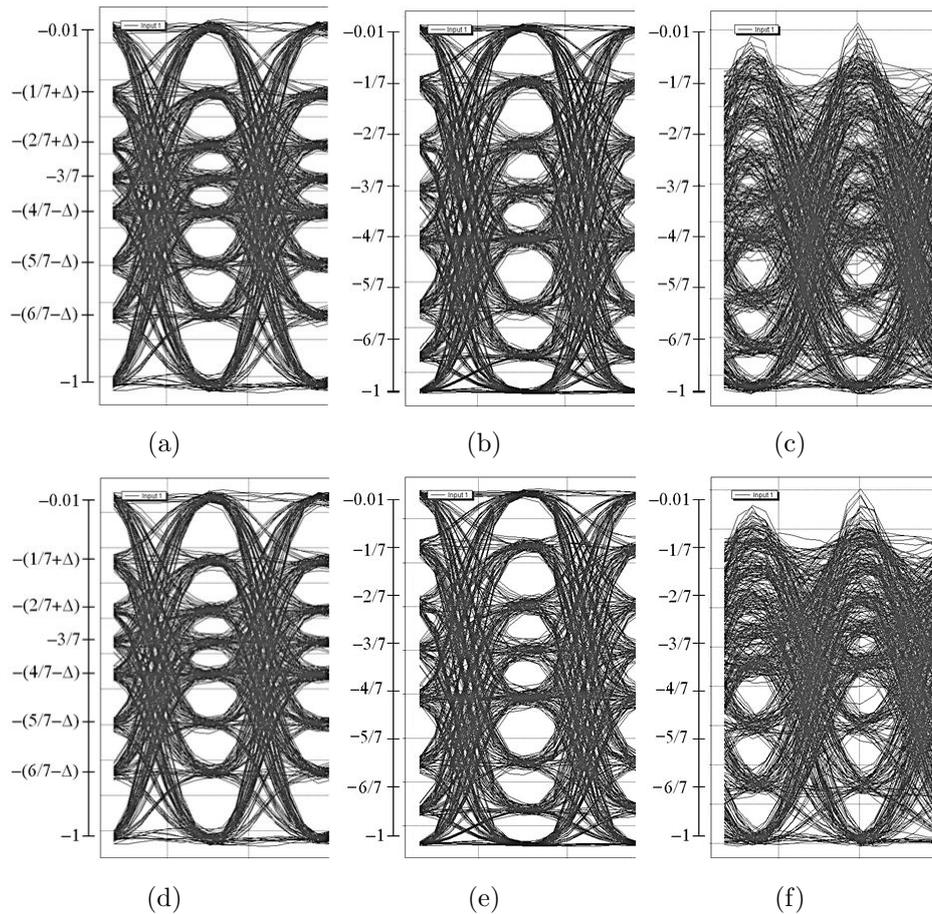


Figure 5.19. Diagramme de l'œil de PAM 8. Les lignes représentent la puissance reçue : -15 dBm ((a), (b) et (c)) et -10 dBm ((d), (e) et (f)). (a) et (d) montrent signal électrique PAM à l'entrée de l'EAM ayant passé par la méthode de compensation de la distorsion. (b) et (e) montrent signal optique PAM à la sortie de l'EAM, en ayant utilisé (a) et (d) en entrée. (c) et (f) montrent signal électrique PAM à la sortie du récepteur SOA-PIN. Paramètres utilisés : $Bias = -0,3 \text{ V}$, $Drive = -1,95 \text{ V}$ et $\Delta = 0,05$, pour les deux puissances reçues.

Δ à tous les niveaux (comme sur la méthode initiale), mais avec des poids différents pour chaque niveau. Ceci permet d'avoir plus de souplesse au niveau de la compensation, sans devoir augmenter les dimensions des balayages et donc le temps de simulation. A priori, ces poids devraient aussi être déterminés en faisant des balayages ; toutefois, nous avons démarré la recherche de ces poids à partir de la Figure 5.19, réduisant ainsi la plage des balayages.

La nouvelle méthode de compensation consiste alors à remplacer les niveaux pré-distordus du signal électrique PAM 8 à l'entrée du modulateur optique du tableau de l'étape 2. de la méthode initiale, par ceux présentés dans le Tableau 5.5. La nouvelle méthode doit être appliquée à chaque valeur de puissance reçue, car les résultats varient en fonction de la saturation de l'amplificateur optique.

La Figure 5.20 présente les diagrammes de l'œil correspondants aux cas précédents (Figure 5.19), une fois appliquée la nouvelle méthode de compensation de la distorsion introduite

Niveaux pré-distordus de PAM 8							
0,01	$1/7+3 \times \Delta$	$2/7+2 \times \Delta$	$3/7$	$4/7-\Delta$	$5/7-2 \times \Delta$	$6/7-3 \times \Delta$	1

Tableau 5.5. Niveaux pré-distordus des signaux PAM 8, pour compenser la distorsion introduite par l'ensemble EAM + SOA-PIN.

par l'ensemble EAM + SOA-PIN. Encore une fois, les lignes correspondent à la puissance optique reçue : -15 dBm pour la 1^{ère} ligne (figures 5.20 (a), (b) et (c)) et -10 dBm pour la 2^{ème} ligne (figures 5.20 (d), (e), et (f)). Les figures 5.20 (a) et (d) montrent le signal électrique PAM à l'entrée du modulateur étant passé par la nouvelle méthode de pré-distorsion. Les figures 5.20 (b) et (e) montrent le signal optique PAM à la sortie de l'EAM, en ayant utilisé (c) et (f) en entrée. Les figures 5.20 (c) et (f) montrent le signal électrique PAM à la sortie du récepteur SOA-PIN.

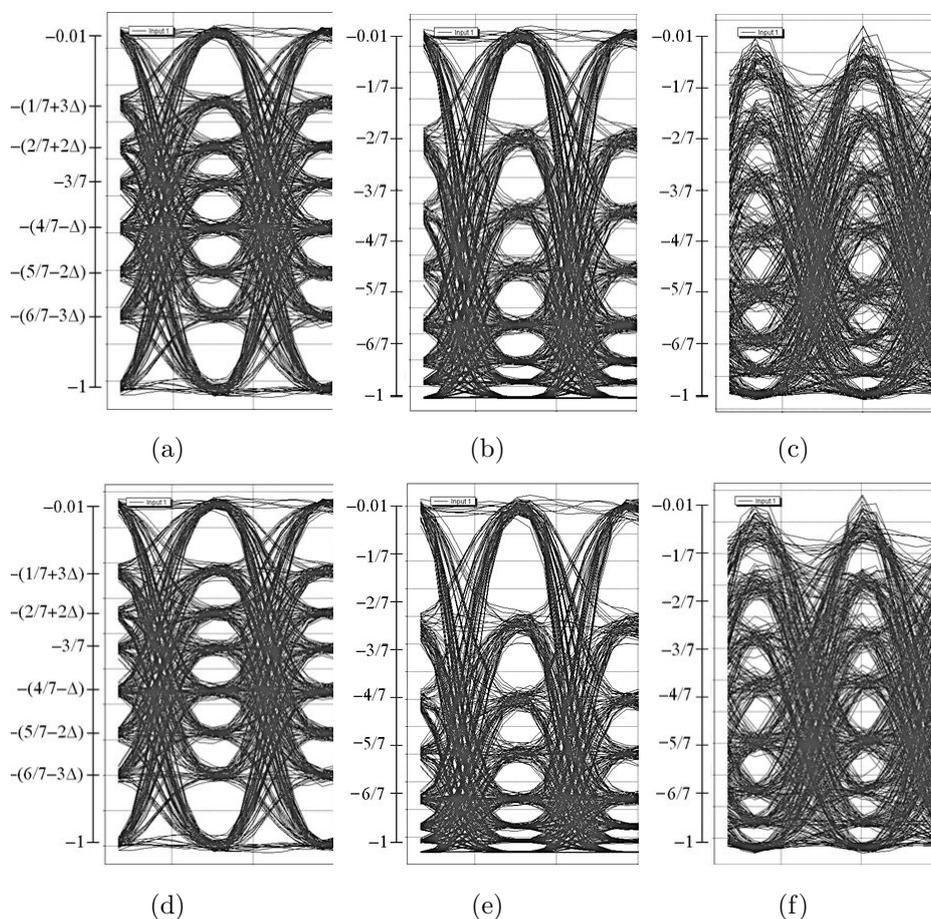


Figure 5.20. Diagramme de l'œil de PAM 8. Les lignes représentent la puissance reçue : -15 dBm ((a), (b) et (c)) et -10 dBm ((d), (e) et (f)). (a) et (d) montrent le signal électrique PAM à l'entrée du modulateur ayant passé par la nouvelle méthode de compensation de la distorsion introduite par l'ensemble EAM + SOA-PIN. (b) et (e) montrent le signal optique PAM à la sortie de l'EAM, en ayant utilisé (a) et (d) en entrée. (c) et (f) montrent le signal électrique PAM à la sortie du récepteur SOA-PIN. Paramètres utilisés : $Bias = -0,7/-0,81$ V, $Drive = -2,1/-1,95$ V et $\Delta = 0,02/0,02$, pour des puissances reçues de -15/-10 dBm.

5.3.3 Résultats de simulation

Nous présentons des résultats de SER en fonction de la puissance reçue pour les formats de modulation PAM 4 et 8 à différents débits. Pour les réseaux POPI et E-POPI nos résultats correspondent aux simulations en accord avec le schéma de la Figure 5.14. Pour le réseau POPI+ nous séparons les structures intra et inter-racks. Pour la structure intra-rack nous considérons, de même, le schéma de simulation de la Figure 5.14. Pour la structure inter-racks, nos résultats sont en accord avec les schémas des simulations des figures 5.15(a) et 5.15(b), qui incluent les éventuelles dégradations causées par un ou deux amplificateurs optiques, respectivement.

Dans ces simulations l'espacement des cinq canaux est de 50 GHz pour les débits de 14, 28, 56 Gbit/s et de 100 GHz pour 112 Gbit/s.

5.3.3.1 POPI et E-POPI

La Figure 5.21 présente des résultats applicables pour les réseaux POPI et E-POPI (aussi valables pour POPI+ intra-rack). La colonne de gauche montre les résultats obtenus avec le format de modulation PAM 4; celle de droite, avec PAM 8. Les lignes 1^{ère}, 2^{ème}, 3^{ème} et 4^{ème} représentent un débit de 14, 28, 56 et 112 Gbit/s, respectivement.

Sur cette figure nous pouvons clairement observer un SER décroissant en fonction de la puissance reçue, qui change d'allure en fonction du récepteur utilisé. Nous observons de même, toujours pour une puissance optique reçue constante, une évolution croissante du SER en fonction du débit pour les deux formats de modulation.

Le SER est plus élevé pour PAM 8. Comme nous l'avons expliqué précédemment, ce format est plus sensible au bruit et aux interférences inter-symboles pour un même taux d'extinction. De plus, nous avons vu que les deux algorithmes de compensation de la distorsion utilisés pour ces simulations sont beaucoup moins précis pour PAM 8 que pour PAM 4, dû à l'utilisation d'une variation commune à tous les niveaux (et non pas adaptée selon le niveau).

Nous considérons toujours un code correcteur d'erreurs binaires qui requiert un BER minimal de 10^{-3} . Encore une fois, le simulateur nous donne les résultats en termes de SER. Alors nous utilisons la limite du HD-FEC à un BER égal au SER afin de représenter le pire des cas possibles. Nous extrayons ainsi, depuis les courbes de SER, les valeurs de puissance reçue pour lesquelles le SER est égal à 10^{-3} (voir le Tableau 5.6). Ces valeurs nous permettront de calculer le nombre maximal de serveurs connectés en utilisant les équations du bilan de liaison de chaque réseau.

5.3.3.2 POPI+

5.3.3.2.1 Intra-rack + Inter-racks (un amplificateur optique)

Nous réduisons le nombre de cas d'études afin de diminuer le temps de simulation. Les

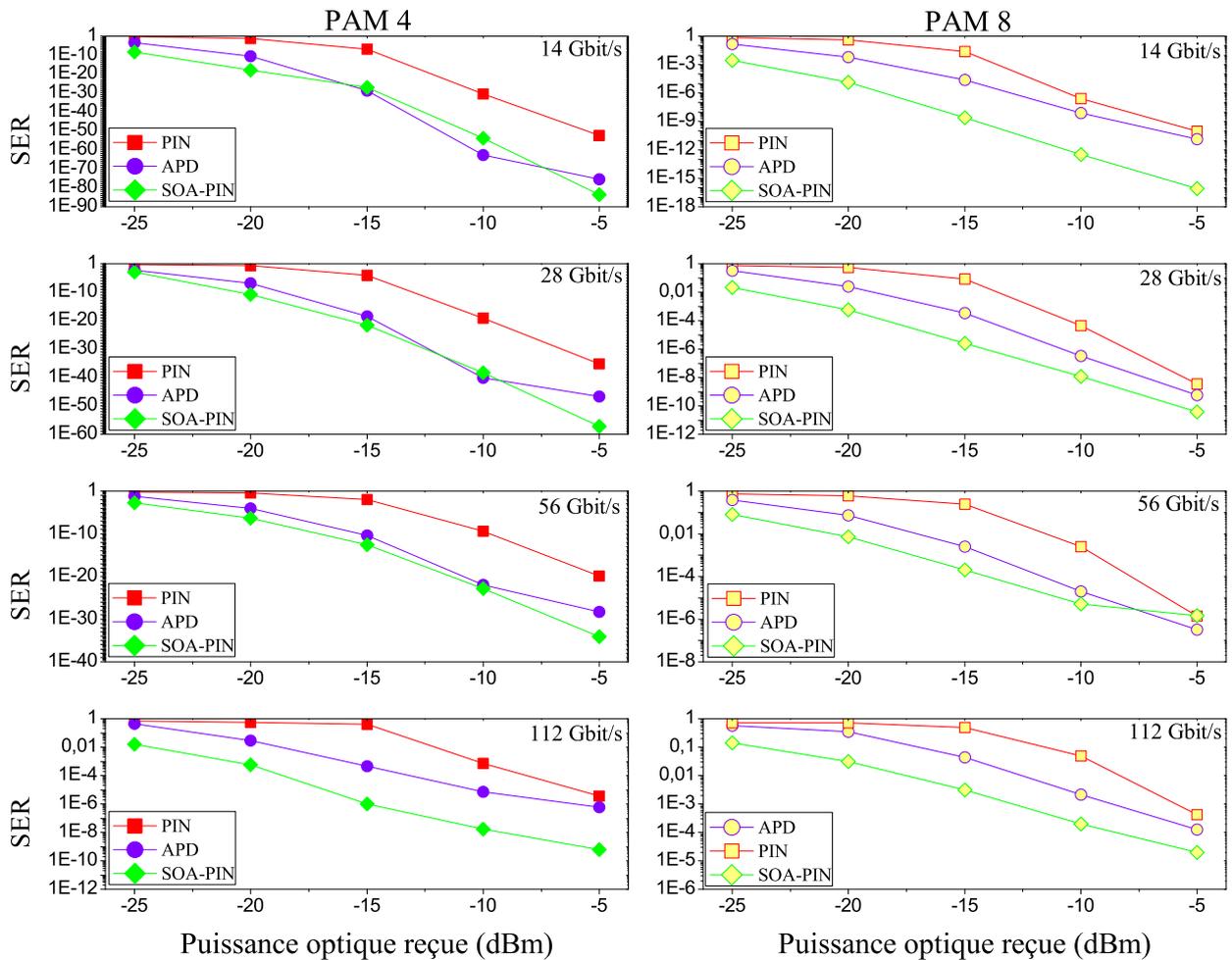


Figure 5.21. SER en fonction de la puissance optique reçue, pour PAM 4 (colonne de gauche) et PAM 8 (colonne de droite), pour différents débits : 14 ((a) et (b)), 28 ((c) et (d)), 56 ((f) et (g)) et 112 Gbit/s ((h) et (i)).

Format de modulation Débit (Gbit/s) \ Récepteur	PAM 4			PAM 8		
	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
14	-18,6	-25,4	-28,1	-13,8	-18,6	-24,5
28	-16,6	-24,2	-25	-12,1	-16,3	-20,7
56	-14,3	-21,9	-24,6	-9,4	-14,1	-17,2
112	-10,3	-16	-21,7	-5,9	-8,7	-13

Tableau 5.6. Puissance optique nécessaire pour un SER de 10^{-3} (dBm), pour les formats de modulation PAM 4 et 8, à des débits de 14, 28, 56 et 112 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN.

cas les plus intéressants à étudier dans cette section correspondent aux trois meilleurs cas de la Figure 5.21 à 112 Gbit/s. Par meilleurs cas nous entendons les cas pour lesquelles un SER constant requiert les puissances reçues les moins élevées.

En conséquence, la Figure 5.22 présente des résultats applicables pour le réseaux POPI+

(un EDFA) à 112 Gbit/s. Les lignes représentent les différents cas choisis : 1^{ère} : PAM 4 avec un récepteur APD ; 2^{ème} : PAM 4 avec un récepteur SOA-PIN ; et 3^{ème} : PAM 8 également avec un récepteur SOA-PIN.

Afin d'estimer les dégradations (pénalités de puissance) apportées par l'EDFA, nous comparons sur la Figure 5.22 les résultats pour les connexions intra (schéma de simulation de la Figure 5.14) et inter-racks (schéma de simulation de la Figure 5.15(a)), respectivement.

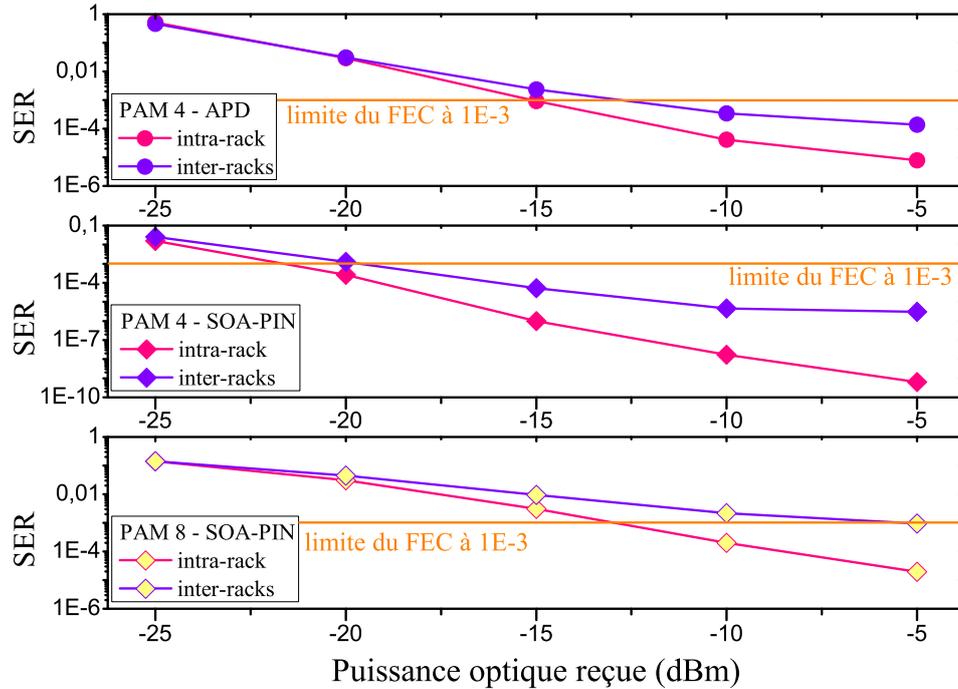


Figure 5.22. SER en fonction de la puissance optique reçue pour les trois meilleurs cas : PAM 4 avec un récepteur APD ; PAM 4 avec un récepteur SOA-PIN ; et PAM 8 avec un récepteur SOA-PIN, pour le réseau POPI+ (un EDFA) à 112 Gbit/s.

De cette figure, nous tirons les valeurs de puissance reçue pour un SER de 10^{-3} et nous calculons les pénalités de puissance dues à l'EDFA, voir Tableau 5.7.

Cas étudié	Puissance optique nécessaire pour un SER de 10^{-3} (dBm)		Pénalités de puissance dues au EDFA (dB)
	Intra-rack	Inter-racks	
PAM 4 - APD	-16	-12,7	3,3
PAM 4 - SOA-PIN	-21,7	-19,7	2
PAM 8 - SOA-PIN	-13	-15	8

Tableau 5.7. Puissance optique nécessaire pour un SER de 10^{-3} pour les connexions intra et inter-racks (un EDFA) du réseau POPI+ et pénalités de puissance introduites par l'EDFA.

Comme prévu, ces pénalités sont plus élevées pour le récepteur APD, qui est moins performant que le SOA-PIN.

Enfin, nous avons évalué l'impact sur le dimensionnement d'une dispersion des paramètres des principaux dispositifs. En effet, il semble indispensable de tenir compte de cette dispersion dans la mesure où nous étudions des systèmes où plusieurs dizaines de composants différents sont susceptibles d'être utilisés. Nous incluons des résultats en faisant varier de ± 5 et de $\pm 10\%$ les paramètres de pré-compensation de la distorsion *Bias*, *Drive* et Δ . De même, nous diminuons le gain et augmentons le facteur de bruit des amplificateurs (EDFA en ligne de transmission et SOA en réception) de 5 et 10%. Ces résultats sont présentés sur la Figure 5.23.

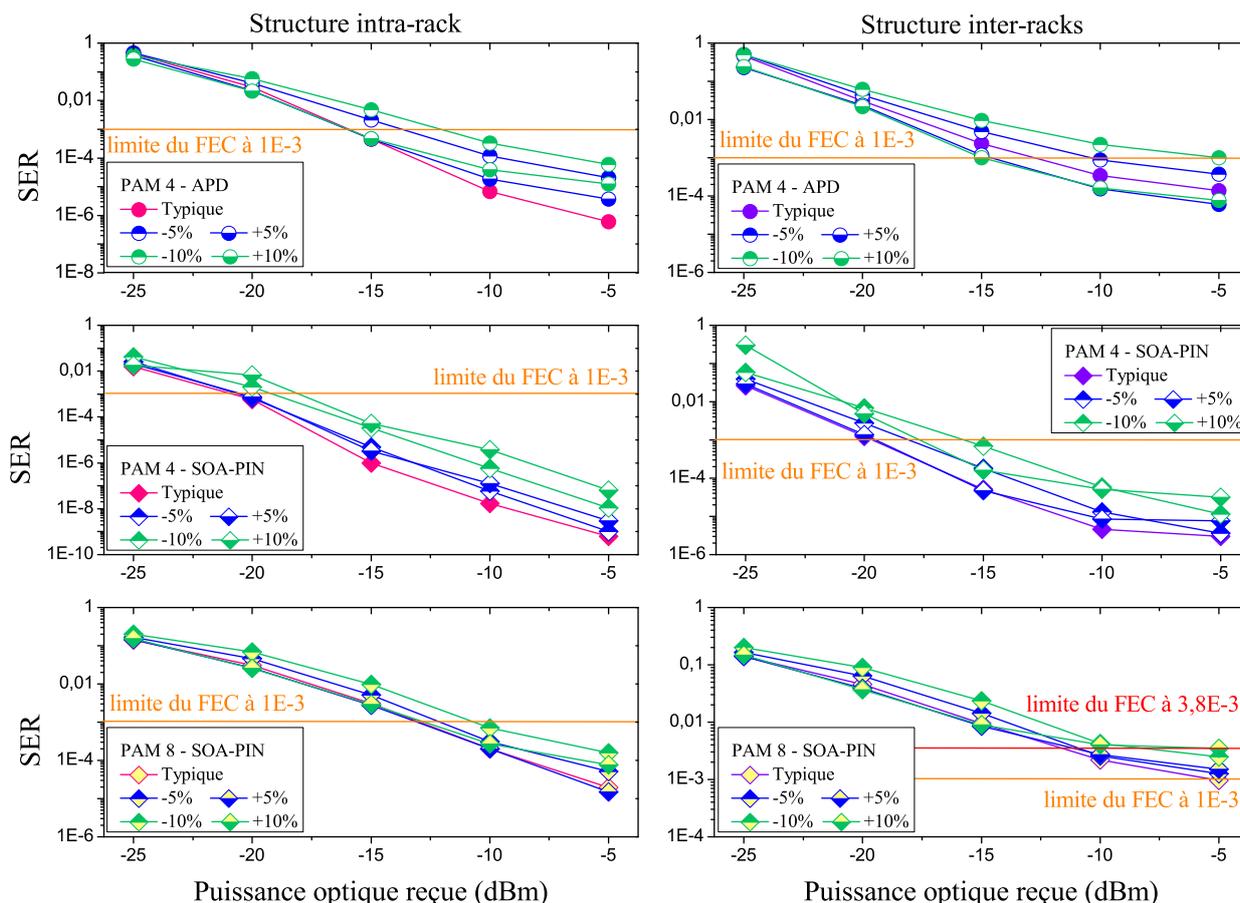


Figure 5.23. SER en fonction de la puissance optique reçue pour les trois meilleurs cas : PAM 4 - APD, PAM 4 - SOA-PIN et PAM 8 - SOA-PIN à 112 Gbit/s, pour les connexions optiques intra et inter-racks du réseau POPI+ (un EDFA). La légende « Typique » représente les résultats sans aucune variation des paramètres de *Bias*, *Drive*, Δ , G_{SOA} , NF_{SOA} , G_{EDFA} ou NF . Les légendes « $\pm 5\%$ » et « $\pm 10\%$ » indiquent que les paramètres de *Bias*, *Drive* et Δ ont été modifiés de ± 5 ou $\pm 10\%$, respectivement. Dans tous les cas (± 5 et $\pm 10\%$) le gain et le facteur de bruit des deux amplificateurs ont été diminué et augmenté de 5 ou 10%.

Nous constatons que, pour certaines de ces courbes (cas PAM 4 - APD et PAM 8 - SOA-PIN), les résultats de + ou - 5% présentent des pénalités de puissance différentes (0 et 2,3 dB pour le cas PAM 4 - APD), et que, les résultats de + ou - 10% peuvent présenter des pénalités de puissance inférieures à celles de $\pm 5\%$ (3,3 et 0 dB). Autrement dit, le fait de

baisser le gain et d'augmenter le bruit est moins pénalisant que le fait de ne pas avoir un paramétrage correct de pré-compensation de la distorsion.

En outre, nous observons que pour le cas PAM 8 - SOA-PIN, tout écart $\geq \pm 5\%$ présente des valeurs de SER supérieures à 10^{-3} pour les connexions inter-racks. Nonobstant, des études récentes montrent que des FEC permettant jusqu'à $3,8 \times 10^{-3}$ de BER sont aujourd'hui réalisables [34],[140]. Ceci permettrait, non seulement d'augmenter la capacité de connexions au réseaux POPI, E-POPI et POPI+, mais améliorerait la robustesse du système en termes de dispersion. Par exemple, le cas PAM 8 - SOA-PIN admettrait jusqu'à $\pm 5\%$ de variations.

Finalement, nous résumons les valeurs de puissance optique nécessaire pour un SER de 10^{-3} dans le Tableau 5.8.

Cas étudié	Intra-rack					Inter-racks				
	Typique	-5%	+5%	-10%	+10%	Typique	-5%	+5%	-10%	+10%
PAM 4 - APD	-16	-13,7	-16	-12,7	-16	-12,7	-10,4	-14,8	-5	-15
PAM 4 - SOA-PIN	-21,7	-20,6	-20,5	-19,2	-18	-19,7	-18,1	-19,6	-15,8	-17,7
PAM 8 - SOA-PIN	-13	-12,4	-13,1	-10,6	-12,8	-5	-2,2	-2	NA	NA

NA : N'est pas applicable, représente les cas où il n'y a pas de $SER \leq 10^{-3}$.

Tableau 5.8. Puissance optique nécessaire pour un SER de 10^{-3} avec une dispersion de ± 5 et $\pm 10\%$, pour les connexions intra et inter-racks (un EDFA) du réseau POPI+.

5.3.3.2.2 Chaîne complète : Intra-rack + Inter-racks (deux amplificateurs optiques)

Nous étudions maintenant la chaîne complète de POPI+ avec les deux amplificateurs optiques en ligne de transmission, à l'aide du schéma de simulation de la Figure 5.15(b). Nous considérons, de même, les trois meilleurs cas à 112 Gbit/s. Encore une fois nous étudions l'effet de la dispersion des paramètres de simulation. Cette fois-ci, nous prenons en compte seulement les pires cas : -5 et -10%. Nous présentons ces résultats sur la Figure 5.24. Nous n'avons pas inclus les résultats du cas PAM 8 - SOA-PIN sur cette figure, parce que toutes les valeurs de SER, sans ou avec dispersion, sont supérieures à 10^{-3} (et même à $3,8 \times 10^{-3}$) pour toutes les valeurs de la puissance optique reçue. Nous résumons les valeurs de puissance optique nécessaire pour un SER de 10^{-3} dans le Tableau 5.9.

Cette fois-ci, nous observons que pour le cas PAM 4 - APD, toute variation du paramétrage présente des valeurs de SER supérieures à 10^{-3} . Cependant, pour une puissance optique reçue de -5 dBm nous retrouvons des valeurs de SER inférieures à $3,8 \times 10^{-3}$.

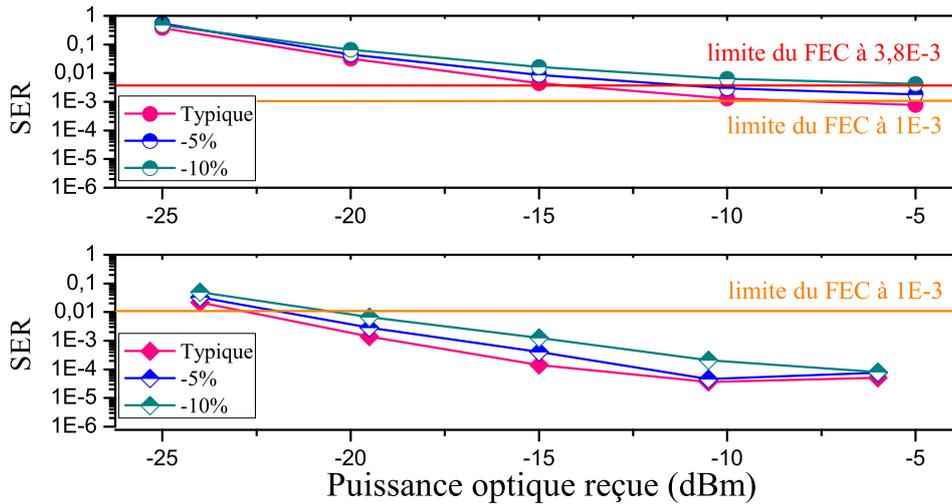


Figure 5.24. SER en fonction de la puissance optique reçue pour les deux meilleurs cas : PAM 4 - APD et PAM 4 - SOA-PIN à 112 Gbit/s, pour les connexions du réseau POPI+ (deux EDFA). La légende « Typique » représente les résultats sans aucune variation des paramètres. Les légendes « -5% » et « -10% » indiquent des variations de -5 et -10% sur les paramètres de *Bias*, *Drive* et Δ , respectivement. Dans tous les cas le gain et le facteur de bruit des deux amplificateurs ont été diminué et augmenté de 5 ou 10%.

Cas étudié	Typique	-5%	-10%
PAM 4 - APD	-7,5	NA ^a	NA ^b
PAM 4 - SOA-PIN	-19,3	-16,3	-14,4

NA : N'est pas applicable, $SER \geq 10^{-3}$.

^a $P_{Rx} = -11,5$ et dBm, pour $SER = 3,8 \times 10^{-3}$.

^b $P_{Rx} = -5$ dBm, pour $SER = 3,8 \times 10^{-3}$.

Tableau 5.9. Puissance optique nécessaire pour un SER de 10^{-3} (dBm) avec une dispersion de -5 et -10%, pour le réseau POPI+, chaîne complète.

5.3.4 Capacité maximale de POPI, E-POPI et POPI+

5.3.4.1 Contraintes imposées par l'émetteur accordable et le filtrage

Non seulement les contraintes physiques, estimées par les bilans de liaison de chaque réseau, mais aussi des contraintes imposées par l'émetteur accordable en longueur d'onde et le filtrage optique, vont déterminer le nombre maximal de serveurs connectés aux réseaux POPI, E-POPI et POPI+.

Le laser accordable en longueur d'onde, que nous considérons tout au long de ce manuscrit [26]-[27], peut configurer jusqu'à 64 canaux espacés de 50 GHz dans une bande de 3200 GHz, avec un temps d'accordabilité inférieur à 50 ns.

Pour accroître autant que possible le nombre de canaux, nous allons considérer des grilles fixes avec un espacement allant de 25 jusqu'à 100 GHz à un pas de 12,5 GHz, en tenant

compte de la bande passante du signal PAM ($1,5 \times$ le débit symbole).

Nous serons donc amenés à considérer des dimensionnements avec des canaux dont l'espacement est inférieur à celui utilisé pour les simulations présentées précédemment. Nous reviendrons sur ce point au vu des résultats de dimensionnement.

Nous calculons maintenant le nombre maximal de serveurs connectés par ces réseaux \hat{N}_{X-GHz} , en utilisant une grille de X GHz. \hat{N}_{X-GHz} ne considère pas les contraintes de la couche de transmission. Nous réservons deux ou trois longueurs d'onde selon le réseau :

- POPI et E-POPI : 2 pour les messages de *report* et *grant* et 1 pour le *gateway*.
- POPI+ intra-rack : 2 pour les messages de *report* et *grant* intra-racks.
- POPI+ inter-racks : 2 pour les messages de *report* et *grant* inter-racks et 1 pour le *gateway*.

Le Tableau 5.10 présente les valeurs de \hat{N}_{X-GHz} pour POPI, E-POPI (1 – 2 – 1 – 1, 1 – 1 – 2 – 1 et 1 – 1 – 1 – 2) et POPI+.

Réseau	POPI	E-POPI	POPI+
$X(GHz)\backslash\hat{N}_{X-GHz}$	$3200/X - 3$	$2 \times 3200/X - 3^*$	$\underbrace{\hat{N}_{X-GHz,intra-rack}}_{3200/X - 2} \times \underbrace{\hat{N}_{X-GHz,inter-racks}}_{3200/X - 3}$
25	125	253	$126 \times 125 = 15750$
37,5	82	167	$83 \times 82 = 6806$
50	61	125	$62 \times 61 = 3906$
62,5	48	99	$49 \times 48 = 2352$
75	39	82	$40 \times 39 = 1560$
87,5	33	70	$34 \times 33 = 1122$
100	29	61	$30 \times 29 = 870$

Tableau 5.10. Nombre maximal de serveurs connectés par les réseaux POPI, E-POPI et POPI+, selon la taille de grille choisie. Seulement les contraintes de l'émetteur accordable en longueur d'onde et l'espacement entre canaux ont été pris en compte.

Concernant le filtrage optique, nous devons vérifier s'il y a des filtres passe-bandes optiques réalisables par rapport à la technologie actuelle. Pour ceci, nous considérons des filtres passe-bande trapézoïdaux ayant une pente P maximale de 800 dB/nm [56].

La Figure 5.25 caractérise le filtre passe-bande trapézoïdal considéré. (a) montre l'allure et les différents paramètres qui caractérisent le filtre : atténuation At (de 20 dB), largeurs de bande minimale B_{min} et maximale B_{max} qui correspondent à la bande passante du signal PAM et à l'espacement des canaux WDM, et la pente P . (b) montre un exemple où le spectre d'un signal PAM 8 est superposé avec un filtre passe-bande trapézoïdal.

Le Tableau 5.11 présente les pentes des filtres passe-bande trapézoïdaux requises pour des signaux PAM 4 et 8, à des débits de 14, 28, 56 et 112 Gbit/s. La couleur verte représente

les pentes inférieures à 800 dB/nm. Les valeurs en couleur noire correspondent aux pentes supérieures à cette valeur. « NA » représente les cas où $B_{max} \leq B_{min}$. En conséquence, allons prendre en compte uniquement les cas représentés en couleur verte pour le dimensionnement de la section suivante.

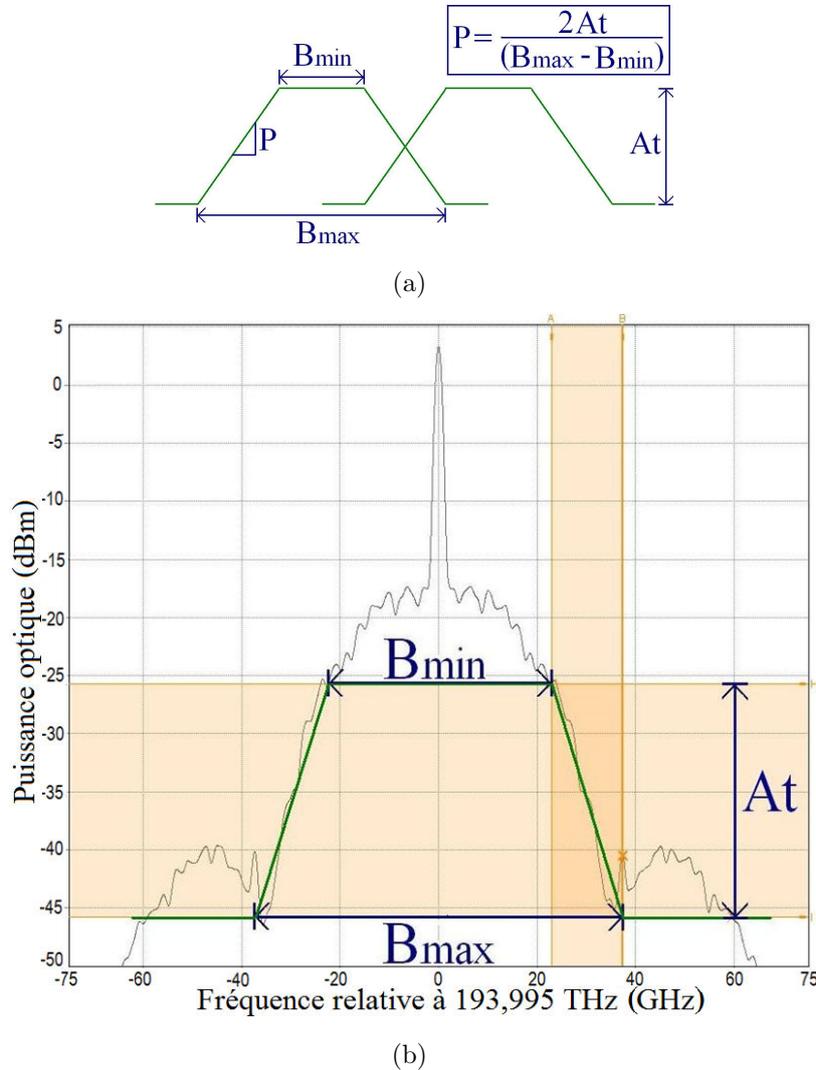


Figure 5.25. Caractérisation des filtres trapézoïdaux. (a) Allure du filtre. At , B_{min} , B_{max} et P correspondent aux paramètres suivants : atténuation, largeurs de bande minimale et maximale et pente, respectivement. (b) Spectre d'un signal PAM 8 superposé avec un filtre trapézoïdal.

5.3.4.2 Contraintes imposées par les pertes d'insertion

5.3.4.2.1 POPI et E-POPI

Pour calculer le nombre maximal de serveurs connectés aux réseaux POPI et E-POPI (et éventuellement de POPI+ intra-rack), en prenant en compte uniquement les contraintes de la couche de transmission, \hat{N}_{IL} (IL : Insertion Losses), nous considérons :

Débit (Gbit/s)	14		28		56		112	
Format de modulation	PAM 4	PAM 8						
B_{max} (GHz) \ B_{min} (GHz)	10,5	4,7	21	14	42	28	84	56
12,5	2500	909	NA	NA	NA	NA	NA	NA
25	345	278	1250	455	NA	NA	NA	NA
37,5	185	164	303	213	NA	526	NA	NA
50	127	116	172	139	625	227	NA	NA
62,5	96	90	120	103	244	145	NA	769
75	78	74	92	82	152	106	NA	263
87,5	65	62	75	68	110	84	1429	159
100	56	54	63	58	86	69	313	114

NA : N'est pas applicable, représente les cas où $B_{max} \leq B_{min}$ et nous ne pouvons pas calculer la pente du filtre.

La couleur verte représente des pentes inférieurs à 800 dB/nm, il s'agit des filtres réalisables selon [56].

Tableau 5.11. Pente du filtre passe-bande optique trapézoïdal (dB/nm) requise pour PAM 4 et 8 à des débits de 14, 28, 56 et 112 Gbit/s.

$$\hat{N}_{IL} = \lfloor 10^{\hat{L}_C/10} \rfloor \quad [\text{partie entière par défaut}] \quad (5.3.2)$$

Où \hat{L}_C représente les pertes maximales du coupleur ($\hat{L}_{C(total)} = \hat{L}_C + \hat{L}_{C,excès}$), qui vont être estimées en accord avec les bilans de liaison de chaque réseau.

Les équations 5.3.3 et 5.3.4 présentent le calcul de \hat{L}_C pour les réseaux POPI et E-POPI (1 – 1 – 2 – 1), respectivement.

$$\hat{L}_C = P_{Tx} - P_{Rx} - L_{EAM} - L_{fibre} - L_{connecteurs} - L_{BPF} - L_{circulateur} - L_{FBG} - M - \hat{L}_{C-excès} \quad (5.3.3)$$

$$\hat{L}_C = P_{Tx} - P_{Rx} - L_{EAM} - L_{fibre} - L_{connecteurs} - L_{BPF} - L_{circulateur} - L_{FBG} - L_{démux} - M - \hat{L}_{C-excès} \quad (5.3.4)$$

Alors, le nombre maximal de serveurs \hat{N} connectés aux réseaux POPI et E-POPI correspond à :

$$\hat{N} = \min\{\hat{N}_{X-GHz}; \hat{N}_{IL}\} \quad (5.3.5)$$

5.3.4.2.2 POPI+

D'autre part, pour le cas particulier de POPI+, nous calculons séparément le nombre maximal de serveurs par rack $\hat{N}_{IL,intra-rack}$ et de racks $\hat{N}_{IL,inter-racks}$ connectés, en prenant en compte uniquement les contraintes de la couche de transmission. Nous estimons donc le

nombre maximal de serveurs connectés à POPI+ dû à ces contraintes comme ceci : $\hat{N}_{IL} = \hat{N}_{IL,intra-rack} \times \hat{N}_{IL,inter-racks}$.

Ainsi, nous utilisons l'équation 5.3.6 pour le dimensionnement de la structure intra-rack de POPI+ (voir schéma de la figure 4.3).

$$\begin{aligned} \hat{L}_{C,intra-rack} = & P_{Tx} - P_{Rx,intra-rack} - L_{EAM} - L_{fibre,intra-rack} - L_{connecteurs} \\ & - L_{BPF} - L_{circulateur} - L_{FBG} - L_{mux} - M_{intra-rack} - \hat{L}_{C,intra-rack-excès} \end{aligned} \quad (5.3.6)$$

Les paramètres P_{Tx} , L_{EAM} , $L_{connecteurs}$, L_{BPF} , $L_{circulateur}$ et L_{FBG} de cette équation prennent les mêmes valeurs que pour POPI et E-POPI. De plus, $P_{Rx,intra-rack} = P_{Rx}$ correspond à la puissance reçue requise pour avoir un SER de 10^{-3} ; $L_{fibre,intra-rack} = L_{fibre} = 0,2$ dB, aux pertes dans la fibre optique (1 km); $L_{mux} = L_{démux} = 1$ dB, aux pertes du multiplexeur en émission; et $M_{intra-rack} = M$, à une marge de 1,5 dB.

Pour le dimensionnement de la structure inter-racks de POPI+ nous utiliserons deux types de simulation. Le premier type prend en compte uniquement l'amplificateur de l'étage inter-racks dans la chaîne de simulation (comme dans la section 5.3.3.2.1). Tandis que le second prend également en compte le *booster* situé dans la passerelle (comme dans la section 5.3.3.2.2). La deuxième méthode est plus exacte, mais la première nous permet de réduire le temps de simulation notamment lorsque nous tenons compte de la dispersion des paramètres.

Le dimensionnement de la structure inter-racks de POPI+ est calculé à l'aide de l'équation suivante :

$$\begin{aligned} \hat{L}_{C,inter-racks} = & P_{in,inter-racks} - P_{Rx,inter-racks} - G_{OA(in-line)} - L_{fibre} - L_{connecteurs} \\ & - L_{WSS} - M_{inter-racks} - \hat{L}_{C,inter-racks-excès} \end{aligned} \quad (5.3.7)$$

Avec $P_{in,inter-racks} = P_{Rx,POPI} - L_{BPF} + G_{OA(booster)}$, où $P_{Rx,POPI}$ correspond à la puissance réelle reçue par la structure intra-rack, qui est calculée à partir de 5.3.8 :

$$\begin{aligned} P_{Rx,POPI} = & P_{Tx} - L_{EAM} - L_{fibre,intra-rack} - L_{connecteurs} - L_{BPF} \\ & - L_{mux} - M_{intra-rack} - L_{C,intra-rack} - L_{C,intra-rack-excès} \end{aligned} \quad (5.3.8)$$

Où $L_{C,intra-rack}$ représente les pertes réelles du coupleur de la structure intra-rack, selon :

$$L_{C,intra-rack} = 10 \times \log \left(\min \{ \hat{N}_{X-GHz,intra-rack}; \hat{N}_{IL,intra-rack} \} \right) \quad (5.3.9)$$

Selon le type de simulation nous utiliserons les valeurs de $P_{Rx,intra-rack}$ tirées du Tableau 5.8 ou du Tableau 5.9.

5.3.4.3 Dimensionnement POPI, E-POPI et POPI+

5.3.4.3.1 POPI et E-POPI

Le nombre maximal \hat{N} de serveurs connectés aux réseaux POPI et E-POPI est calculé pour les formats de modulation PAM 4 et 8 à des débits de 14, 28, 56 et 112 Gbit/s. Ces résultats sont présentés dans l'Annexe B.

Nous observons que pour les débits plus faibles (14 et 28 Gbit/s) la contrainte prédominante est notamment celle de la plage d'accord et la largeur du canal. Autrement dit, nous pourrions encore diminuer la puissance en émission (<15 dBm, jusqu'à que les deux limitations soient égales) dans le but d'économiser de l'énergie.

Lorsque nous augmentons le débit la contrainte de pertes d'insertion devient plus importante et nous commençons à avoir besoin d'augmenter la puissance en émission pour les récepteurs moins efficaces (PIN et APD) et les formats de modulation plus sensibles au bruit (PAM 8).

Les graphiques présentés sur la Figure 5.26 résument les résultats obtenus pour les meilleurs cas de POPI et E-POPI.

Sur la Figure 5.26(b) nous avons inclus des résultats pour une puissance en émission de 25 mW. En effet, le réseau E-POPI a environ le double de pertes d'insertion au niveau du coupleur par rapport au réseau POPI, il est donc nécessaire d'augmenter la puissance en émission pour relâcher cette contrainte et atteindre le maximum d'entités connectées.

Pour les réseaux POPI et E-POPI le maximum d'entités connectées est atteint avec le format PAM 8 en utilisant le récepteur SOA-PIN.

Le Tableau 5.12 montre l'espacement entre canaux nécessaire pour les meilleurs cas étudiés. Nous notons que pour certains de ces cas, notamment pour les débits de 14 et 28 Gbit/s, les espacements requis sont inférieurs à ceux que nous avons pris en compte pour obtenir les sensibilités des récepteurs en simulation. Ceci semble être contraignant pour le dimensionnement que nous venons de présenter. Cependant, en ce qui concerne POPI, le nombre maximal de serveurs connectés calculé est limité principalement par la plage d'accord et la largeur du canal et non pas par les pertes d'insertion ($\hat{N}_{X-GHz} \ll \hat{N}_{X-GHz}$). Autrement dit, les résultats obtenus ne changeraient pas.

Débit (Gbit/s)	PAM 4-APD $P_{TX} = 15$ dBm	PAM 4-SOA-PIN $P_{TX} = 15$ dBm	PAM 8-SOA-PIN $P_{TX} = 20$ ou 25 dBm
14	25	25	25
28	37,5	37,5	25
56	50	50	37,5
112	100	100	62,5

Tableau 5.12. Espacement entre canaux requis (GHz) pour les meilleurs cas étudiés.

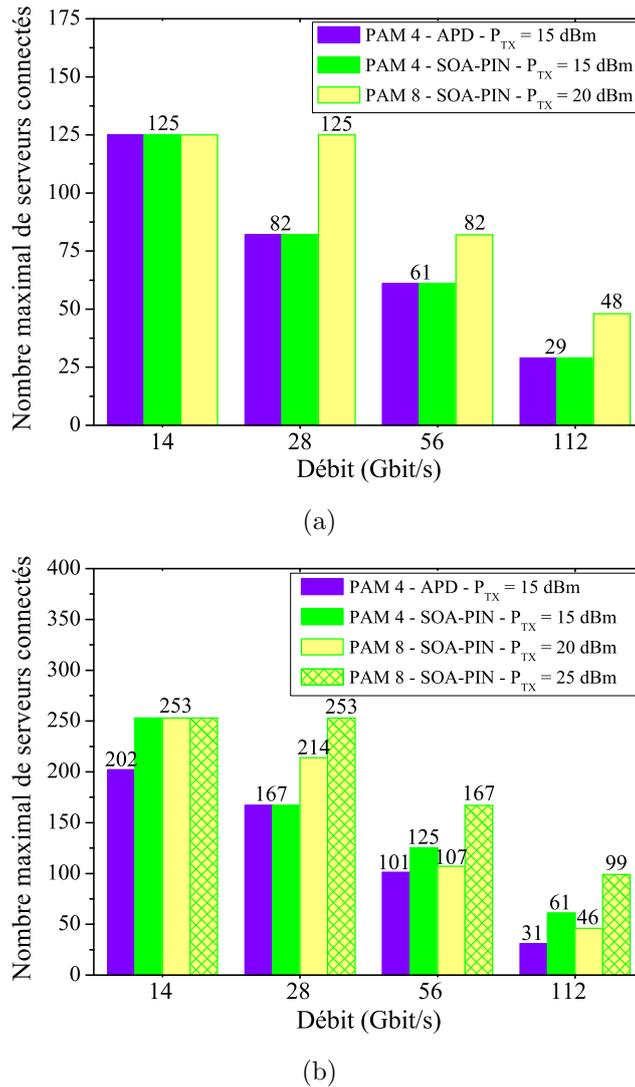


Figure 5.26. Dimensionnement des réseaux (a) POPI et (b) E-POPI.

5.3.4.3.2 POPI+

Le nombre maximal de racks connectés au réseau POPI+ (en ne considérant que le bruit d'émission spontanée amplifiée d'un seul EDFA) $\hat{N}_{inter-racks}$ est calculé pour les trois meilleurs cas étudiés à 112 Gbit/s. (A) PAM 4 - APD, $P_{Tx} = 15$ dBm ; (B) PAM 4 - SOA-PIN, $P_{Tx} = 15$ dBm ; (C) PAM 8 - SOA-PIN, $P_{Tx} = 20$ dBm et (D) PAM 4 - SOA-PIN, $P_{Tx} = 25$ dBm. Comme pour ces deux derniers cas la contrainte de pertes insertion est la plus importante, nous estimons intéressant d'étudier un cas additionnel (E), en augmentant la puissance émise à 30 dBm. Certes travailler à des puissances aussi élevées (classe 4) augmente la consommation d'énergie, les risques de dégradation des connecteurs et de la fibre optique et la dangerosité oculaire. Cependant, ces calculs nous permettent d'envisager un dimensionnement qui pourrait être possible, par exemple en augmentant le gain des SOA en pré-amplification à la réception à 25 ou 30 dB. Il faudrait bien sûr évaluer l'influences du bruit ASE. Ces résultats sont présentés dans le Tableau 5.13. Encore une fois, les couleurs

bleue et rouge représentent la contrainte du réseau la plus importante.

Dimensionnement de POPI+ à 112 Gbit/s					
Cas	A	B	C	D	E
$\hat{N}_{intra-rack}$	30	30	46	49	49
$L_{C,intra-rack}$ (dB)	17,7	17,7	19,9	20,2	20,2
$P_{Rx,POPI}$ (dBm)	-15,7	-15,7	-12,9	-8,2	-3,2
$P_{in,inter-racks}$ (dBm)	-0,7	-0,7	2,1	6,8	11,8
$P_{Rx,inter-racks}$ (dBm)	-12,7	-19,7	-5	-5	-5
$\hat{L}_{C,inter-racks}$ (dB)	16,1	22,1	11,9	15,9	20,2
$\hat{N}_{IL,inter-racks}$	37	158	12	35	102
$\hat{N}_{inter-racks}$	29	29	12	35	48

Cas A : PAM 4 - APD, espacement de 100 GHz, $P_{Tx} = 15$ dBm
 Cas B : PAM 4 - SOA-PIN, espacement de 100 GHz, $P_{Tx} = 15$ dBm
 Cas C : PAM 8 - SOA-PIN, espacement de 62,5 GHz, $P_{Tx} = 20$ dBm
 Cas D : PAM 8 - SOA-PIN, espacement de 62,5 GHz, $P_{Tx} = 25$ dBm
 Cas E : PAM 8 - SOA-PIN, espacement de 62,5 GHz, $P_{Tx} = 30$ dBm

Tableau 5.13. Nombre maximal de serveurs par rack et de racks connectés aux étages intra-rack et inter-racks, respectivement, du réseau POPI+, avec -5 et -10% de dispersion sur les paramètres de simulation.

Le nombre maximal de serveurs par rack et de racks connectés au réseau POPI+ (en ne considérant que le bruit de bruit d'émission spontanée amplifiée d'un seul EDFA) \hat{N} et donc calculé pour chaque cas en considérant une dispersion sur les paramètres de simulation de -5 et -10%, voir Tableau 5.14.

Le nombre maximal total de serveurs connectés à POPI+ pour chaque cas est représenté sur la Figure 5.27.

Nous constatons que le cas B est le plus robuste, car il est le seul à supporter une dispersion de -10% sans diminuer le nombre maximal total de serveurs connectés à POPI+.

Dans le cas E le maximum de machines connectées atteint 2352, cependant, ce cas ne tolère que jusqu'à -5% de dispersion sur les paramètres de simulation.

Finalement, nous vérifions le dimensionnement de POPI+ en considérant les effets de deux amplificateurs optiques en ligne de transmission. Les résultats sont présentés dans le Tableau 5.15.

Nous constatons par ce tableau que le récepteur APD ne tolère pas le bruit d'émission spontanée amplifiée de deux amplificateurs optiques à 112 Gbit/s. Pourtant, le cas B se montre toujours robuste en gardant un \hat{N} de $30 \times 29 = 870$.

Cette étude montre que le débit de 112 Gbit/s est encore trop ambitieux pour le réseau POPI+ en utilisant des systèmes de transmission avec des récepteurs moins efficaces que le

Dimensionnement de POPI+ à 112 Gbit/s avec -5 et -10% de dispersion			
Cas	Typique	-5%	-10%
A	30 serveurs × 29 racks	20 serveurs × 29 racks	16 serveurs × 13 racks
B	30 serveurs × 29 racks	30 serveurs × 29 racks	30 serveurs × 29 racks
C	46 serveurs × 12 racks	41 serveurs × 7 racks	NA
D	49 serveurs × 35 racks	49 serveurs × 19 racks	NA
E	49 serveurs × 48 racks	49 serveurs × 48 racks	NA

Cas A : PAM 4 - APD, espacement de 100 GHz, $P_{Tx} = 15$ dBm

Cas B : PAM 4 - SOA-PIN, espacement de 100 GHz, $P_{Tx} = 15$ dBm

Cas C : PAM 8 - SOA-PIN, espacement de 62,5 GHz, $P_{Tx} = 20$ dBm

Cas D : PAM 8 - SOA-PIN, espacement de 62,5 GHz, $P_{Tx} = 25$ dBm

Cas E : PAM 8 - SOA-PIN, espacement de 62,5 GHz, $P_{Tx} = 30$ dBm

NA : N'est pas applicable. Représente les cas pour lesquels $\forall P_{Tx}, SER > 10^{-3}$, pour le réseau inter-racks.

Tableau 5.14. Nombre maximal de serveurs par rack et de racks connectés aux étages intra-rack et inter-racks, respectivement, du réseau POPI+.

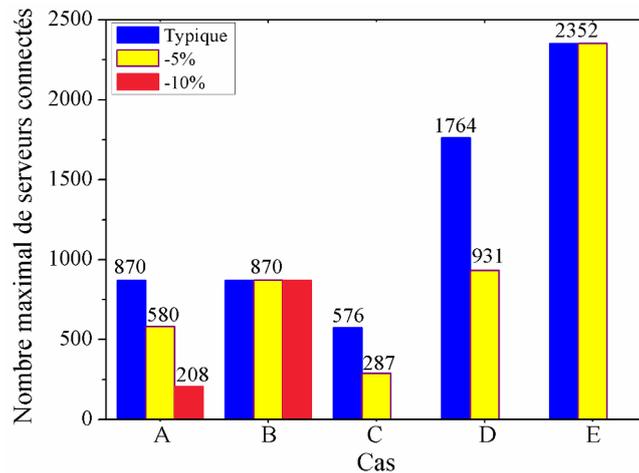


Figure 5.27. Dimensionnement du réseau POPI+ avec -5 et -10% de dispersion sur les paramètres de simulation.

SOA-PIN ou des formats de modulation multi-niveaux au-dessus PAM 4. Cependant, vu le comportement de ce format en fonction du débit pour les réseaux POPI et E-POPI, il serait intéressant d'analyser le dimensionnement de POPI+ à un débit plus faible, soit 56 Gbit/s. Ceci non seulement relâcherait la contrainte de pertes d'insertion, mais aussi celle de la plage d'accord et de la largeur du canal dû à sa bande passante inférieure.

5.4 Conclusion

Dans ce chapitre nous avons étudié le format de modulation d'amplitude à M états de façon expérimentale et par simulation par ordinateur. Nous avons mis en place une chaîne

Dimensionnement de POPI+ à 112 Gbit/s avec -5 et -10% de dispersion			
Cas	Typique	-5%	-10%
A	29 serveurs × 12 racks	NA	NA
B	30 serveurs × 29 racks	30 serveurs × 29 racks	30 serveurs × 29 racks

Cas A : PAM 4 - APD, espacement de 100 GHz, $P_{Tx} = 15$ dBm
Cas B : PAM 4 - SOA-PIN, espacement de 100 GHz, $P_{Tx} = 15$ dBm

NA : N'est pas applicable. Représente les cas pour lesquels $\forall P_{Tx}, SER > 10^{-3}$, pour le réseau inter-racks.

Tableau 5.15. Nombre maximal de serveurs par rack et de racks connectés aux étages intra-rack et inter-racks, respectivement, du réseau POPI+ (le bruit d'émission spontanée amplifiée des deux EDFA est considéré).

de transmission optique avec les formats de modulation PAM 4 et 8, avec un débit maximal de 18 Gbit/s. Nous avons utilisé comme récepteur une photodiode PIN sans et avec pré-amplification optique à l'aide d'un SOA. Nous avons comparé et validé nos résultats expérimentaux avec des résultats de simulation.

Nous avons ensuite étudié les mêmes formats de modulation à des débits plus élevés (jusqu'à 112 Gbit/s) par simulation par ordinateur. Nous avons étudié les récepteurs PIN, APD et SOA-PIN. Nous avons appliqué une méthode de pré-compensation de la distorsion introduite par le modulateur optique et par l'ensemble EAM + SOA quand nous utilisons le récepteur SOA-PIN.

Grâce aux résultats obtenus par simulation nous avons calculé le bilan de liaison pour les architectures d'interconnexion POPI, E-POPI et POPI+. Ces calculs nous ont permis d'évaluer les contraintes du dimensionnement de ces réseaux dues aux pertes d'insertion de leurs composants.

Nous avons, en outre, analysé les contraintes de dimensionnement dues aux caractéristiques du laser accordable utilisé par ces réseaux, ainsi que par l'espacement entre canaux.

Nous avons ainsi estimé la capacité maximale des réseaux POPI, E-POPI et POPI+, avec le format de modulation PAM 4 et 8, à des débits de 14, 28, 56 et 112 Gbit/s, pour les trois récepteurs étudiés. Le nombre maximal de serveurs connectés à 112 Gbit/s est de 48, 99 et 2352 serveurs, pour POPI, E-POPI et POPI+, respectivement. Ces résultats ont été obtenus avec PAM 8, avec le récepteur SOA-PIN.

Nous avons ensuite tenu compte d'une possible dispersion des paramètres des composants. Les résultats obtenus montrent que pour les réseaux POPI et E-POPI le nombre maximal de serveurs connectés reste constant. Cependant, pour le réseau POPI+ le format de modulation PAM 8 supporte jusqu'à ± 5 % de dispersion. Ainsi que PAM 4 se montre robuste même devant $\pm 10\%$ de dispersion. La capacité maximale de POPI+ est donc réduite à 870 serveurs.

Une des difficultés rencontrées tient au fait que les méthodes de pré-compensation utilisées manquent de flexibilité dans le choix de leurs paramètres de correction pour ne pas

trop alourdir les simulations. Il serait donc intéressant d'évaluer une solution hybride qui consisterait à utiliser la méthode déjà proposé avec les poids selon le Tableau 5.5 comme valeurs initiales des nouvelles variations $\Delta_{0,1,2,\dots,7}$ ($\Delta_0 = 0$, $\Delta_1 = 3 \times \Delta$, $\Delta_2 = 3 \times \Delta$, $\Delta_3 = 0$, $\Delta_4 = -\Delta$, $\Delta_5 = -2 \times \Delta$, $\Delta_6 = -3 \times \Delta$). Le fait de démarrer la recherche des $\Delta_{0,1,2,\dots,8}$ avec ces valeurs initiales devrait réduire considérablement le nombre d'itérations à faire avec des balayages en 8 dimensions.

CONCLUSIONS ET PERSPECTIVES

6.1 Conclusions

Dans ce travail nous nous sommes intéressés aux interconnexions optiques des centres de données qui sont des éléments fondamentaux des réseaux de télécommunications modernes. Nous avons plus particulièrement intéressés, dans le cadre du projet EPOC, aux DC de petite et moyenne taille qui utilisent des énergies renouvelables. Ces DC ont besoin d'un réseau d'interconnexion à haut débit qui puisse assurer une migration rapide des VM des serveurs, tout en étant efficace énergétiquement. Ce qui nous a conduit à privilégier des solutions à base des paquets optiques pour concilier haut débit et faible consommation.

Nous rappelons ci-dessous nos principales contributions à l'étude des réseaux de paquets optiques dédiés aux connexions internes des centres des données.

Nous avons choisi de prendre comme base le réseau d'interconnexion POPI parce qu'il repose sur une infrastructure entièrement passive. Nous avons également décidé de mettre l'accent sur des formats de modulations multiniveaux utilisant une détection directe (PAM 4 et 8) de façon à augmenter l'efficacité spectrale ou le débit, tout en limitant les coûts des interfaces.

Une première évaluation de la performance de POPI en termes de latence et débit a été réalisée par les auteurs qui ont proposé ce réseau. Nos études sur POPI se sont concentrées sur sa couche physique.

Nous avons d'abord étudié le bilan de liaison de la version initiale de POPI et nous avons proposé trois variantes qui limitent le nombre de connexions directes au coupleur central, permettant de réduire les pertes d'insertion d'environ 3 dB.

Une autre limitation est due au fait que POPI utilise des sources accordables en longueur d'onde. Nous pouvons augmenter le nombre de canaux en travaillant sur deux axes : en utilisant une ou plusieurs bandes de transmission optiques, ou en adaptant au mieux possible l'espacement entre canaux avec la largeur spectrale du signal selon le format de modulation utilisé. Le choix se fera également en fonction des technologies disponibles et des nouvelles grilles flexibles de multiplexage optique WDM.

Le premier axe nous a conduit à proposer des extensions de POPI sous le nom d'E-POPI.

Il s'agit de trois extensions qui permettent de doubler le nombre d'entités connectées en utilisant deux émetteurs accordables dans des bandes de transmission optiques différentes. Nous avons proposé et analysé les composants optiques à utiliser pour chaque extension. Enfin, nous avons comparé ces trois extensions notamment en termes de pertes d'insertion et de consommation d'énergie. Ces options peuvent en principe être étendues à plusieurs bandes de transmission.

Nous avons ensuite proposé une nouvelle architecture hiérarchique à deux étages sous le nom de POPI+, qui permet d'augmenter au carré le nombre d'entités connectées par rapport au réseau POPI originel. Le premier étage consiste en un réseau POPI conventionnel qui sera utilisé uniquement pour les transferts intra-racks. Nous introduisons un nouveau réseau qui sera utilisé exclusivement pour des transferts inter-racks. Ce dernier est basé sur un coupleur $n : 1$, un amplificateur optique et des WSS. Chacun des réseaux opère dans une bande différente.

Nous avons étudié particulièrement la structure de contrôle de cette nouvelle architecture. Une unité centrale de contrôle va gérer les transmissions inter-racks en collaboration avec les contrôleurs locaux de chaque réseau POPI intra-rack. Nous détaillons la structure du réseau qui connecte l'unité centrale de contrôle au contrôleurs locaux et nous évaluons également ses pertes d'insertion.

La structure précédente est adaptée au cas où le trafic intra-rack est dominant. Dans le cas contraire, nous proposons une variante de POPI+ sous le nom de T-POPI qui facilite les transferts inter-racks grâce à l'utilisation d'une bande de transmission différente à celle des transferts internes au rack de destination. Nous analysons de même les pertes d'insertion de cette architecture.

Pour atteindre des débits au-delà de 10 Gbit/s, nous avons étudié le format de modulation PAM 4 et 8 niveaux. Le bilan de liaison ainsi que les contraintes dues au spectre disponible sont pris en compte pour dimensionner les architectures d'interconnexion POPI, E-POPI et POPI+. Cette étude a été faite à plusieurs débits, avec différents types de récepteurs (PIN, APD et SOA-PIN). Nos principaux résultats ont été obtenus par simulation par ordinateur. Ce simulateur a été validé expérimentalement à plus bas débit dans le cas des récepteurs PIN et SOA-PIN.

Nous avons utilisé de la pré-compensation de la distorsion adaptée aux non-linéarités du le modulateur électro-absorbant et à l'ensemble de ce modulateur plus le récepteur SOA-PIN. L'utilisation de cette pré-compensation nous a permis d'améliorer les sensibilités des récepteurs tout en gardant des temps de simulation raisonnables.

Le nombre maximal de machines connectées au réseaux POPI, E-POPI et POPI+ est de 48, 99 et 2352, respectivement, avec le format de modulation PAM 8 en utilisant le récepteur SOA-PIN.

Finalement, nous avons tenu compte d'une dispersion de 5 et 10% sur les paramètres de

l'algorithme de compensation de la distorsion ainsi que sur le gain et le facteur de bruit des amplificateurs à fibre et à semi-conducteur. Les résultats obtenus ont montré que le format de modulation PAM 8, qui maximisait le nombre d'entités connectées, ne peut supporter que jusqu'à 5% de dispersion pour POPI+. Le format PAM 4 quant à lui, se montre toujours robuste en présence de 10% de dispersion pour les trois réseaux étudiés, ainsi le nombre maximal de machines connectées à POPI+ est de 870.

Nous avons pu ainsi proposer des solutions d'interconnexion à base de paquets optiques pour les centres de données comportant quelques dizaines de serveurs, tels que ceux considérés par le projet EPOC, susceptibles de concilier faible consommation d'énergie et débits élevés jusqu'à 100 Gbit/s. Nous avons aussi introduit des extensions des architectures précédentes à des centres de données de taille plus importante, comportant de l'ordre du millier de serveurs.

6.2 Perspectives

Dans un premier temps, nous proposons d'approfondir nos travaux dans l'aspect de couche de transmission optique avec les perspectives suivantes :

- Améliorer les méthodes de la pré-compensation de la distorsion comme proposé dans le Chapitre 5.
- Prendre en compte des gains du SOA plus élevés (jusqu'à 35 dB), comme présenté récemment dans l'état de l'art [44].
- Valider le simulateur optique à l'aide des manipulations expérimentales à des débits plus élevés.

Ensuite, nous proposons de changer certaines hypothèses que nous avons prises en compte, notamment sur le format de modulation et la structure de contrôle de nos réseaux.

- Lorsque les technologies permettant de faire une détection cohérente deviennent plus simples et économiques, étudier des formats type QAM ou CAP dans le contexte des DC afin d'améliorer l'efficacité spectrale par rapport au format PAM.
- Pour POPI+ proposer une structure de contrôle local indépendante du réseau des données (par exemple, avec un réseau ethernet). Ceci relâcherait les contraintes de pertes d'insertion. Chaque rack gagnerait au moins deux serveurs (car nous libérons les connexions des contrôleurs au coupleur central de POPI) et le bilan de liaison économiserait 1 dB de pertes (de l'ensemble FBG et circulateur). Ceci permettrait d'augmenter le nombre de racks connectés (le nombre de serveurs par rack n'augmenterait pas, puisque la contrainte de plage d'accordabilité et d'espacement entre canaux est la plus forte pour les connexions intra-racks).

Finalement, il serait utile d'évaluer la performance de bout en bout des réseaux POPI, E-POPI et POPI+ avec leur systèmes respectifs de contrôle, pour des débits de 14, 28, 56 et 112 Gbit/s.

BIBLIOGRAPHIE

- [1] C. Kachris and I. Tomkos, “Optical interconnection networks for data centers”, IEEE Proceedings of 17th International Conference on Optical Network Design and Modeling, ONDM 2013, 16-19 Apr. 2013, Brest, France.
- [2] J. Koomey, “Growth in Datacenter Electricity Use 2005 to 2010”, Analytics Press, Oakland, CA, USA, August 2011.
- [3] C. Kachris and I. Tomkos, “Power Consumption Evaluation of All-optical Data Center Networks”, IEEE Cluster Computing, The Journal of Networks, Software Tools and Applications, vol. 16, issue 3, 2012.
- [4] J. Sinsky and P. Winzer, “A Microwave Engineer’s Perspective 100-Gb/s”, IEEE Microwave Magazine, Digital Object Identifier 10.1109/MMM.2008.931669, Apr. 2009.
- [5] Y. Chen, J. Li, P. Zhu, B. Guo, L. Zhu, Y. He and Z. Chen, “Experimental demonstration of 400 Gb/s optical PDM-OFDM superchannel multicasting by multiple-pump FWM in HNLF”, Optics Express, vol. 21, issue 8, pp. 9915-9922, 2013.
- [6] X. Zhou and L. Nelson, “Advanced DSP for 400 Gb/s and Beyond Optical Networks”, IEEE Journal of Lightwave Technology, vol. 32, no. 16, pp. 2716-2725, Aug. 2014.
- [7] K. Kechaou, « Modélisation et caractérisation de sources optiques pour les réseaux d’accès et métropolitains », thèse pour obtenir le grade de docteur délivré par Télécom ParisTech, Paris, 13 Déc. 2012.
- [8] “Alcatel 1905 LMI Up to 30mW WDM C-band version for external modulation CW 1.55 μ m Laser Module with optical Isolator”, Mar. 2012, disponible sur : www.3spgroup.com/data/File/3S\%20PHOTONICS/Produits/3SP_1905_LMI_CBand.pdf, consulté le 23 novembre 2017.
- [9] H. Khaleghi, « Influence des amplificateurs optiques à semi-conducteurs (SOA) sur la transmission cohérente de signaux optiques à format de modulation multi-porteuses (CO-OFDM) », thèse pour obtenir le grade de docteur de l’université de Bretagne occidentale, Brest, 30 Nov. 2012.
- [10] S. Murata, I. Mito and K. Kobayashi, “Over 720 GHz (5.8 nm) Frequency Tuning by a 1.5 μ m DBR Laser with Phase and Bragg Wavelength Control Regions”, Electronic Letters, vol. 23, no. 8, pp. 403-405, Apr. 1987.

- [11] F. Delorme, A. Ramdane, B. Rose, S. Slemptes and H. Nakajima, "Ultra-Fast Optical Switching Operation of DBR Lasers Using an Electro-Optical Tuning Section", *IEEE Photonics Technology Letters*, vol. 7, no. 3, pp. 269-271, Mar. 1995.
- [12] F. Delorme, S. Slemptes, G. Alibert, B. Rose and J. Brandon, "Butt-jointed DBR Laser with 15 nm Tunability Grown in Three MOVPE Steps", *Electronic Letters*, vol. 31, no. 15, pp. 1244-1245, Jul. 1987.
- [13] Y. Tohmori, Y. Yoshikuni, T. Tamamura, M. Yamamoto, Y. Kondo and H. Ishii, "Ultra-wide Wavelength Tuning with Single Longitudinal Mode by Super Structure Grating (SSG) DBR Lasers", *Conference Digest. 13th IEEE International Semiconductor Laser Conference 1992*, 21-25 Sept. 1992, Kagazwa, Japan.
- [14] Y. Tohmori, Y. Yoshikuni, H. Ishii, F. Kano, T. Tamamura, Y. Kondo and M. Yamamoto, "Broad-Range Wavelength-Tunable Superstructure Grating (SSG) DBR Lasers", *IEEE Journal of Quantum Electronics*, vol. 29, no. 6, pp. 1817-1823, Jun. 1993.
- [15] H. Ishii, F. Kano, Y. Tohmori, Y. Kondo, T. Tamamura and Y. Yoshikuni, "Broad Range (34nm) Quasi-Continuous Wavelength Tuning in Super-Structure-Grating DBR Lasers", *Electronic Letters*, vol. 30, no. 14, pp. 1134-1135, Jul. 1994.
- [16] M. Oberg, P-J. Rigole, S. Nilsson, T. Klinga, L. Backbom, K. Streubel, J. Wallin and T. Kjellberg, "Complete Single Mode Wavelength Coverage Over 40 nm with a Super Structure Grating DBR Laser", *IEEE Journal of Lightwave Technology*, vol. 13, no. 9, pp. 1892-1898, Sep. 1995.
- [17] N. Fujiwara, R. Yoshimura, K. Kato, H. Ishii, F. Kano, Y. Kawaguchi, Y. Kondo, K. Ohbayashi and H. Oohashi, "140-nm Quasi-Continuous Fast Sweep Using SSG-DBR Lasers", *IEEE Photonics Technology Letters*, vol. 20, no. 12, pp. 1015-1017, Jun. 2008.
- [18] N. Fujiwara, R. Yoshimura, K. Kato, H. Ishii, Y. Kawaguchi, Y. Kondo and H. Oohashi, "70 nm Fast Wavelength Sweep Using SSG-DBR Lasers", *Conference on Lasers and Electro-Optics - Pacific Rim, 2007. CLEO/Pacific Rim 2007*, 26-31 Aug. 2007, Seoul, South Korea.
- [19] C.K. Gardiner, R.G.S. Plumb, P.J. Williams and T.J. Reid, "Wavelength Tuning in Three Section Sampled Grating DBR Lasers", *Electronic Letters*, vol. 31, no. 15, pp. 1258-1260, Jul. 1995.
- [20] C. Ougier, A. Talneau, F. Delorme, Y. Raffle, S. Slemptes and D. Mathoorasing, "80 Addressable Wavelength Channels over 33 nm with a Sampled-Grating DBR Laser for 2.5 Gb/s WDM Applications", *IEEE Proceedings of 22nd European Conference on Optical Communication ECOC 2012*, 19-23 Sept. 1996, Oslo, Norway.
- [21] F. Delorme, G. Alibert, C. Ougier, S. Slemptes and H. Nakajima, "Sampled-Grating DBR Lasers with 101 Wavelengths over 44 nm and Optimised Power Variation for WDM Applications", *Electronic Letters*, vol. 34, no. 3, pp. 279-281, Feb. 1998.
- [22] B. Mason, G.A. Fish, S.P. DenBaars and L.A. Coldren, "Widely Tunable Sampled Grating DBR Laser with Integrated Electroabsorption Modulator", *IEEE Photonics Technology Letters*, vol. 11, no. 6, pp. 638-640, Jun. 1999.

- [23] B. Mason, J. Barton, G.A. Fish, L.A. Coldren and S.P. DenBaars, "Design of Sampled Grating DBR Lasers with Integrated Semiconductor Optical Amplifiers", *IEEE Photonics Technology Letters*, vol. 12, no. 7, pp. 762-764, Jul. 2000.
- [24] Y.A. Akulova, C. Schow, A. Karim, S. Nakagawa, P. Kozodoy, G.A. Fish, J. DeFranco, A. Dahl, M. Larson, T. Wipiejewski, D. Pavinski, T. Butrie and L.A. Coldren, "Widely-Tunable Electroabsorption-Modulated Sampled Grating DBR Laser Integrated With Semiconductor Optical Amplifier", *Optical Fiber Communication Conference and Exhibit 2002, OFC 2002*, 17-22 Mar. 2002, Anaheim, CA, USA.
- [25] M.C. Larson, M. Bai, D. Bingo, N. Ramdas, S. Penniman, G.A. Fish and L.A. Coldren, "Mode Control of Widely-Tunable SG-DBR Lasers", *IEEE Proceedings of 28th European Conference on Optical Communication ECOC 2002*, 8-12 Sept. 2002, Copenhagen, Denmark.
- [26] J. E. Simsarian, A. Bhardwaj, J. Gripp, K. Sherman, Y. Su, C. Webb, L. Zhang and M. Zirngibl, "Fast-Switching Characteristics of a Widely-Tunable-Laser Transmitter", *IEEE Photonics Technology Letters*, vol. 15, no. 8, pp. 1038-1040, Aug. 2003.
- [27] J. E. Simsarian and L. Zhang, "Wavelength Locking a Fast-Switching Tunable Laser", *IEEE Photonics Technology Letters*, vol. 16, no. 7, pp. 1745-1747, Jul. 2004.
- [28] J.E. Simsarian, M.C. Larson, H.E. Garrett, Hong Xu and T.A. Strand, "Less than 5-ns Wavelength Switching with an SG-DBR Laser", *IEEE Photonics Technology Letters*, vol. 18, no. 4, pp. 565-567, Feb. 2006.
- [29] M. Gotoda, T. Nishimura and Y. Tokuda, "Widely Tunable SOA-Integrated DBR Laser with Combination of Sampled-Grating and Superstructure Grating", *Conference Digest. 2004 IEEE 19th Semiconductor Laser Conference 2004*, 21-25 Sept. 2004, Matsue-shi, Japan.
- [30] R. Radouani, « Dérive dans les modulateurs électro-optiques Mach-Zehnder, analyse physique et résolution », thèse pour obtenir le grade de docteur de l'université Paul Verlaine, Metz, 22 Jun. 2006.
- [31] P. Pecci, « Conception, fabrication et analyse d'une source intégrée laser-modulateur électro-absorbant à ondes progressives pour des transmissions optiques haut débit à 1,55 μm sur InP », thèse pour obtenir le grade de docteur de l'université Lille 1, Lille, 14 Déc. 2000.
- [32] N. Dupuis, « Intégration par épitaxie sélective d'un modulateur électro-absorbant et d'un amplificateur optique à base de puits quantiques AlGaInAs », thèse pour obtenir le grade de docteur de l'université Metz 1, Metz, 21 Jan. 2009.
- [33] P. Angelini, F. Blache, C. Caillaud, P. Charbonnier, B. Duval, M. Goix, F. Jorge, F. Mallecot, K. Mekhazni, J. Y. Dupuy, and M. Achouche, "64 Gb/s Optical Transmission Using DFB-EAM Transmitter and SOA-PIN-TIA Receiver with -23.5 dBm Record Sensitivity", *IEEE Proceedings of 42nd European Conference on Optical Communication ECOC 2012*, 18-22 Sept. 2012, Dusseldorf, Germany.
- [34] M. A. Mestre, H. Mardoyan, C. Caillaud, R. Ríos-Müller, J. Renaudier, P. Jennevé, F. Blache, F. Pommereau, J. Decobert, F. Jorge, P. Charbonnier, A. Konczykowska, J-Y. Du-

- puy, K. Mekhazni, J.-F. Paret, M. Faugeron, F. Mallecot, M. Achouche and S. Bigo, "Compact InP-based DFB-EAM Enabling PAM-4 112 Gb/s Transmission over 2 km", *Journal of Lightwave Technology*, vol. 34, no. 7, pp. 1572-1578, Apr. 2016.
- [35] "HUBER+SUHNER Cube Optics AG", 2017, blog disponible sur : <http://blog.cubeoptics.com/index.php/2015/09/6-bands-you-need-to-know>, consulté le 23 novembre 2017.
- [36] G. P. Agrawal, "Fiber-Optic-Communication Systems", Wiley-Interscience, A John Wiley & Sons INC. Publication, Third Edition, 2002.
- [37] R. J. Mears, L. Reekie, I. M. Jauncey and D. N. Payne, "Low-Noise-Erbium-Doped Fiber Operating at 1.54 μm ", *IEEE Photonics Technology Letters*, vol. 23, no. 19, pp. 1026-1028, 1987.
- [38] E. Desurvire, J. R. Simpson and P.C. Becker, "High-Gain-Erbium-Doped Travelling-Wave-Fiber Amplifier", *Optics Letters*, vol. 12, no. 11, pp. 888-890, Mar. 1987.
- [39] P. Zaca-Morán, E. Kuzin, J. Torres-Turiján, J.G. Ortega-Mendoza, F. Chávez, G.F. Pérez-Sánchez, L.C. Gómez-Pavón, "High Gain Pulsed Erbium-Doped Fiber Amplifier for the nNonlinear Characterization of SWCNTs Photodeposited on Optical Fibers", *Optics & Laser Technology*, vol. 52, pp. 15-20, Jan. 2013.
- [40] D. Aguiar, G. Grasso, A. Righetti and F. Meli, "EDFA with Continuous Amplification of C and L Bands for Submarine Applications", *Microwave and Optoelectronics Conference, IMOC 2015, Porto de Galinhas, Brazil, 3-6 Nov. 2015*.
- [41] Y. Ernori and S. Namiki, "100nm-Bandwidth-Flat-Gain Raman Amplifiers Pumped and Gain-Equalized by 12-Wavelength-Channel WDM High Power Laser Diodes", *Optical Fiber Communication Conference, Optical Society of America, OFC 1999, 21-26 Feb. 1999, San Diego, CA, USA*.
- [42] M. M. J. Martini, C. E. S. Castellani, M. J. Pontes; M. R. N. Ribeiro and H. J. Kalinowski, "Multi-Pump Optimization for Raman+EDFA Hybrid Amplifiers under Pump Residual Recycling", *Microwave and Optoelectronics Conference, IMOC 2009, Belem Do Para, Brazil, 3-6 Nov. 2009*.
- [43] C. Rivera, R. Olivares, M. T. M. Rocco Giraldo, A. P. Lopez Barbero, M. J. Pontes, M. E. V. Segatto, M. A. G. Martinez and J. C. W. Costa, "Numerical Simulations and Experimental Results of a Hybrid EDFA-Raman Amplifier", *Microwave and Optoelectronics Conference, IMOC 2009, Belem Do Para, Brazil, 3-6 Nov. 2009*.
- [44] J. Renaudier, A. Carbo Meseguer, A. Ghazisaeidi, P. Tran, R. Rios Muller, R. Brenot, A. Verdier, F. Blache, K. Mekhazni, B. Duval, H. Debregeas, M. Achouche, A. Boutin, F. Morin, L. Letteron, N. Fontaine, Y. Frignac, G. Charlet, "First 100-nm Continuous-Band WDM Transmission System with 115Tb/s Transport over 100km Using Novel Ultra-Wideband Semiconductor Optical Amplifiers ", *IEEE Proceedings of 43rd European Conference on Optical Communication ECOC 2017, 17-21 Sep. 2017, Gothenburg, Denmark*.

- [45] P. Angelini, « Contribution à l'étude de nouvelles technologies de co-packaging et de co-design appliquées à la réalisation de modules photorécepteurs pour les systèmes de télécommunications de prochaine génération », thèse pour obtenir le grade de docteur de l'université de Limoges, Limoges, 5 Sep. 2017.
- [46] C. Caillaud, G. Glastre, F. Lelarge, R. Brenot, S. Bellini, J-F. Paret, O. Drisse, D. Carpentier and M. Achouche, "Monolithic Integration of a Semiconductor-Optical Amplifier and High-Speed Photodiode with Low Polarization Dependence Loss", *IEEE Photonics Technology Letters*, vol. 24, no. 11, pp. 897-899, Jun. 2012.
- [47] M. Anagnosti, C. Caillaud, J-F. Paret, F. Pommereau, G. Glastre, F. Blache and M. Achouche, "Record Gain x Bandwidth (6.1 THz) Monolithically Integrated SOA Photoreceiver for 100-Gbit/s Applications", invited paper, *Journal of Lightwave Technology*, vol. 33, no. 6, pp. 1186-1190, Mar. 2015.
- [48] C. Caillaud, P. Chanclou, F. Blache P. Angelini, B. Duval, P. Charbonnier, D. Lanteri, G. Glastre and M. Achouche, "Integrated SOA-PIN Detector for High-Speed Short Reach Applications", invited paper, *Journal of Lightwave Technology*, vol. 33, no. 8, pp. 1596-1601, Apr. 2015.
- [49] J. Y. Huh, J. K. Lee, S-K. Kang, J. C. Lee, "Analysis of PAM-N (N=4, 5, 6, 7 and 8) signals operating at 103.125 Gbps for next-generation Ethernet", *12th International Conference on Optical Internet 2014, COIN 2014*, 27-29 Aug. 2014, Jeju, South Korea.
- [50] G. Charlet, J. Renaudier, H. Mardoyan, P. Tran, O. Bertran Pardo, F. Verluise, M. Achouche, A. Boutin, F. Blache, J-Y. Dupuy and S. Bigo, "Transmission of 16.4-bit/s Capacity Over 2550 km Using PDM QPSK Modulation Format and Coherent Receiver", *Journal of Lightwave Technology*, vol. 33, no. 8, pp. 153-157, Feb. 2009.
- [51] S. Kleis and C. G. Schaffer, "Coexistence of PSK quantum key distribution and WDM channels in optical transmission systems", *IEEE Proceedings of 17th ITG-Symposium*, 12-13 May 2016, Leipzig, Germany.
- [52] J. Zhao and A.D. Ellis, "Spectral Efficiency Enhancement Using Coherent WDM with Multi-Level Offset-QAM", *IEEE Proceedings of 37th European Conference on Optical Communication, ECOC 2011*, 18-22 Sep. 2011, Geneva, Switzerland.
- [53] E. Giacomidis, J. L. Wei, X. Q. Jin and J.M. Tang, "Improved Transmission Performance of Adaptively Modulated Optical OFDM Signals over Directly Modulated DFB Laser-based IMDD Links Using Adaptive Cyclic Prefix", *Optics Express*, vol. 16, issue 13, pp. 9480-9494, 2008.
- [54] A. Hamié, M. Hamze, J. L. Wei, A. Sharaiha and J. M. Tang, "Theoretical Investigations of Quantum-dot Semiconductor Optical Amplifier Enabled Intensity Modulation of Adaptively Modulated Optical OFDM Signals in IMDD PON Systems", *Optics Express*, vol. 19, issue 25, pp. 25696-25711, 2011.
- [55] E. Pincemin, "Challenges of 40/100 Gbps Deployments in Long-haul Transport Networks on Existing Fibre and System Infrastructure", *Optical Fiber Communication*, collocated

- National Fiber Optic Engineers Conference 2010, OFC/NFOEC 2010, 21-25 March 2010, San Diego, CA, USA.
- [56] E. Pincemin, M. Song, J. Karaki, O. Zia-Chahabi, T. Guillosoy, D. Grot, G. Thouenon, C. Betoule, R. Clavier, A. Poudoulec, M. Van der Keur, Y. Jaouen, R. Le Bidan, T. Le Gall, P. Gravey, M. Morvan, B. Dumas Feris, M-L. Moulinard and G. Froc, "Multi-band-OFDM Transmission at 100 Gbps with Sub-band-Optical Switching", invited paper, *Journal of Lightwave Technology*, vol. 32, no. 12, pp. 2202-2219, May 2014.
- [57] ITU-T, Terms and definitions for sub-lambda photonic switched networks. International Telecommunication Union-Telecommunication, 2012, Geneva, Switzerland.
- [58] R. Penty, J. D. Ingham, A. Wonfor, K. Wang, I. White, "Reducing the Energy Consumption of Photonics Hardware in Data Center Networks", in the "Energy Efficient Networking for Data Centers" Workshop at Super Computing Conference 2011.
- [59] J. L. Wei, J. D. Ingham, D. G. Cunningham, R. V. Penty and I. H. White, "Performance and Power Dissipation Comparisons between 28 Gb/s NRZ, PAM, CAP and Optical OFDM Systems for Data Communication Applications", *IEEE Journal of Lightwave Technology*, vol. 30, no. 20, pp. 3273-3280, 2012.
- [60] J. L. Wei, J. D. Ingham, D. G. Cunningham, R. V. Penty and I. H. White, "Comparisons between 28 Gb/s NRZ, PAM, CAP and optical OFDM systems for Datacommunication Applications", *IEEE Optical Interconnects Conference 2012, OIC 2012*, 20-23 May 2012, Santa Fe, NM, USA.
- [61] J. D. Ingham, R. V. Penty, I. H. White and D. G. Cunningham, "Carrierless Amplitude and Phase Modulation for Low-Cost, High-Spectral-Efficiency Optical Datacommunication Links", *Conference on Lasers and Electro-Optics (CLEO) and Quantum Electronics and Laser Science Conference (QELS) 2010*, 16-21 May 2010, San Jose, CA, USA.
- [62] J. L. Wei, D. G. Cunningham, R. V. Penty and I. H. White, "Feasibility of 100G Ethernet enabled by Carrierless Amplitude/Phase Modulation and Optical OFDM", *IEEE Proceedings of 38th European Conference on Optical Communication ECOC 2012*, 16-20 Sept. 2012, Amsterdam, Netherlands.
- [63] M. Iglesias Olmedo, T. Zuo, J. Bevensen Jensen, Q. Zhong, X. Xu, S. Popov, and I. Tafur Monroy, "Multiband Carrierless Amplitude Phase Modulation for High Capacity Optical Data Links", *IEEE Journal of Lightwave Technology*, vol. 32, no. 4, pp. 798-804, Feb. 2014.
- [64] S. Qian, X. Luo and Y. Jin, "Low Power Consumption Routing and Spectrum Allocation in Optical OFDM Networks", *8th International Conference on Information, Communications and Signal Processing 2011, ICICS 2011*, 13-16 Dec. 2011, Singapore, Singapore.
- [65] C. Kachris and I. Tomkos, "Energy-Efficient Bandwidth Allocation in Optical OFDM-Based Data Center Networks", *Optical Fiber Communication Conference and Exposition 2012, OFC/NFOEC 2012*, and the *National Fiber Optic Engineers Conference*, 4-8 Mar. 2012, Los Angeles, CA, USA.

- [66] C. Kachris et al., “Energy-Efficient Flexible-Bandwidth OFDM-Based Data Center Network”, IEEE 1st International Conference on Cloud Networking Conference, CLOUDNET 2012, 28-30 Nov. 2012, Paris, France.
- [67] J. L. Wei, D. G. Cunningham, R. V. Penty and I. H. White, “Study of 100 Gigabit Ethernet Using Carrierless Amplitude/Phase Modulation and Optical OFDM”, IEEE Journal of Lightwave Technology, vol. 31, no. 9, pp. 1367-1373, May 2013.
- [68] J. L. Wei, Q. Cheng, R. V. Penty, and I. H. White and D. G. Cunningham, “Analysis of Complexity and Power Consumption in DSP-Based Optical Modulation Formats”, Signal Processing in Photonic Communications 2014, 13-17 Jul. 2014, San Diego, CA, USA.
- [69] SMART 2020 : Enabling the low carbon economy in the information age. A report by The Climate Group on behalf of the Global eSustainability Initiative (GeSI), 2008.
- [70] M. K. Weldon, “The Future X Network : A Bell Labs Perspective”, CRC Press, 2016.
- [71] M. A. Mestre Adrover, « Réseaux optiques pour les centres de données : solutions à court et long terme », thèse pour obtenir le grade de docteur de Télécom SudParis, Paris, 21 Oct. 2016.
- [72] Cisco, “Cisco Global Cloud Index : Forecast and Methodology, 2015-2020”, 2016. Document disponible sur : http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns1175/Cloud_Index_White_Paper.html#wp9000816, consulté le 29 octobre 2017.
- [73] C. Kachris and I. Tomkos, “A Survey on Optical Interconnects for Data Centers”, IEEE Communications Surveys and Tutorials, vol. 14, no. 4, fourth quarter 2012.
- [74] P. Pepeljugoski, J. Kash, F. Doany, D. Kuchta, L. Schares, C. Schow, M. Taubenblatt, B. J. Offrein and A. Benner, “Low Power and High Density Optical Interconnects for Future Supercomputers”, Optical Fiber Communication Conference collocated National Fiber Optic Engineers Conference 2010, OFC/NFOEC 2010, 21-25 Mar. 2010, San Diego, CA, USA.
- [75] M. A. Taubenblatt, “Optical interconnects for high performance computing”, Journal of Lightwave Technology, vol. 30, no. 4, pp. 448-457, Feb. 2012.
- [76] A. Singla, A. Singhz, K. Ramachandran, L. Xuz and Y. Zhang, “Proteus : a Topology Malleable Data Center Network”, IEEE Proceedings of 9th ACM SIGCOMM Workshop on Hot Topics in Networks, ser. Hotnets '10, 2010.
- [77] N. Farrington, G. Porter, S. Radhakrishnan, H. Hajabdolali Bazzaz, V. Subramanya, Y. Fainman, G. Papen and A. Vahdat, “Helios : a Hybrid Electrical/Optical Switch Architecture for Modular Data Centers”, IEEE Proceedings of ACM SIGCOMM 2010 conference on SIGCOMM ser. SIGCOMM'10, 2010.
- [78] X. Ye, P. Mejia, Y. Yin, R. Proietti, S. J. B. Yoo and V. Akella, “DOS : A Scalable Optical Switch for Datacenters”, IEEE Proceedings of 6th ACM Symposium on Architectures for Networking and Communications Systems, ser. ANCS '10, 2010.

- [79] K. Xia, Y-H. Kaob, M. Yangb and H. J. Chaoa, “Petabit Optical Switch for Data Center Networks”, Technical report, Polytechnic Institute of NYU, 2010.
- [80] J. Gripp, J. E. Simsarian, J. D. LeGrange, P. Bernasconi and D. T. Neilson, “Photonic Terabit Routers : The IRIS Project”, Optical Fiber Communication Conference, Optical Society of America, 2010.
- [81] R. Luijten, W. E. Denzel, R. R. Grzybowski and R. Hemenway, “Optical Interconnection Networks : The OSMOSIS project”, 17th Annual Meeting of the IEEE Lasers and Electro-Optics Society, 2004.
- [82] M. A. Mestre, G. de Valicourt, P. Jennev e, H. Mardoyan, S. Bigo and Y. Pointurier, “Optical Slot Switching-based Datacenter with Elastic Burst-Mode Coherent Transponders”, IEEE Proceedings of 40th European Conference on Optical Communication, ECOC 2014, 21-25 Sep. 2014, Cannes, France.
- [83] Y. Gong, Y. Lu, X. Hong, S. He and J. Chen, “Passive Optical Interconnects at Top of the Rack for Data Center Networks”, IEEE Proceedings of 18th Conference on Optical Network Modelling and Design, ONDM 2014, 19-22 May, Stockholm, Sweden 2014.
- [84] J. Chen, Y. Gong, M. Fiorani and S. Aleksic, “Optical Interconnects at the Top of the Rack for Energy-Efficient Data Centers”, IEEE Communications Magazine, vol. 53, no. 8, pp. 140-148, Aug. 2015.
- [85] M. Fiorani, S. Aleksic, M. Casoni, L. Wosinska and J. Chen, “Energy-Efficient Elastic Optical Interconnect Architecture for Data Centers”, IEEE Communications Letters, vol. 18, no. 9, pp. 1531-1534, Sep. 2014.
- [86] R. Indre, J. Pesic and J. Roberts, “POPI : A Passive Optical Pod Interconnect for High Performance Data Centres”, IEEE Proceedings of 18th Conference on Optical Network Modelling and Design, ONDM 2014, 19-22 May, Stockholm, Sweden 2014.
- [87] “Expanding the Role of 3D MEMS Technology to Meet Growing Data Bandwidth Demands”, Calient white paper, 2013, disponible sur : <http://www.calient.net/wp-content/uploads/downloads/2013/08/CALIENT-3D-MEMS-Whitepaper-August-2013-Final.pdf>, consult e le 23 novembre 2017.
- [88] “SDN Optical Transport and DCI Networks”, disponible sur : <http://www.calient.net/solutions/optimized-sdn-transport-and-dci-networks/>, consult e le 23 novembre 2017.
- [89] A. Gumaste and B. M. Krishna Beri, “On The Architectural Considerations of the FIS-SION (Flexible Interconnection of Scalable Systems Integrated using Optical Networks) Framework for Data-Centres”, IEEE Proceedings of 17th Conference on Optical Network Design and Modeling, ONDM 2013, 16-19 Apr. 2013, Brest, France.
- [90] S. J. Ben Yoo, “Optical Packet and Burst Switching Technologies for the Future Optical Internet”, IEEE Journal of Lightwave Technology, vol. 24, no. 12, 2006.

- [91] D. Chiaroni, "Optical Packet Add/drop Multiplexers for Packet Ring Networks", IEEE Proceedings of 34th European Conference on Optical Communication, ECOC 2014, 21-25 Sep. 2008, Brussels, Belgium.
- [92] I. Saniee and I. Widjaja, "A new Optical Network Architecture that Exploits Joint Time and Wavelength Interleaving", Optical Fiber Communication Conference, Optical Society of America, OFC 2004, 9-14 Mar. 2014, San Francisco, California, USA.
- [93] A. Triki, P. Gavignet, B. Arzur, E. Le Rouzic and A. Gravey, "Efficient Control Plane for Passive Optical Burst Switching Network", International Conference on Information Networking 2013, ICOIN 2013, 28-30 Jan. 2013, Bangkok, Thailand.
- [94] N. Farrington, presentation of "Helios : a Hybrid Electrical/Optical Switch Architecture for Modular Data Centers", at IEEE Proceedings of ACM SIGCOMM 2010 conference on SIGCOMM ser. SIGCOMM'10, 2010.
- [95] H. J. Chao, Z. Jing and K. Deng, "PetaStar : A petabit photonic packet switch", IEEE J. Sel. Areas Commun. , Special Issue on High-Performance Optical/Electronic Switches/Routers for High-Speed Internet, vol. 21, pp. 1096-1112, 2003.
- [96] T. Benson, A. Akella and D. Maltz. "Network-Traffic Characteristics of Data Centers in the Wild", Proceedings of 10th ACM SIGCOMM conference on Internet measurement, IMC'10, pp. 267-280, 2010.
- [97] T. Benson, A. Anand, A. Akella and M. Zhang. "Understanding Data-Center-Traffic Characteristics", SIGCOMM Computer Communication Review, vol. 40, no. 1, pp. 92-99, 2010.
- [98] Logiciel « Omnet++ » disponible sur : <https://omnetpp.org/>, consulté le 23 novembre 2017.
- [99] Coupleur passif, documentation disponible sur : <https://www.amphenol-airlb.de/load.php?KID=91>, consultée le 23 novembre 2017.
- [100] Coupleur passif, documentation disponible sur : <https://thorbroadcast.com/datasheets>, consultée le 23 novembre 2017.
- [101] Coupleur passif, documentation disponible sur : <http://www.adcfiber.com/show.asp?nid=136>, consultée le 23 novembre 2017.
- [102] P. Anslow and M. Hajduczenia, "Single-Mode-Fibre Loss", IEEE 802.3av, Orlando, March 2007.
- [103] Fibre optique monomode, information disponible sur : <https://www.cisco.com/c/en/us/support/docs/optical/synchronous-digital-hierarchy-sdh/29000-db-29000.html>, consulté le 23 novembre 2017.
- [104] Connecteurs, documentation disponible sur : http://www.ps-center.co.jp/pdf/conec/fiberoptic_connectors_catalog.pdf, consultée le 23 novembre 2017.
- [105] D. Feng, W. Qian, H. Liang, Ch. Kung, Z. Zhou, Z. Li, J. S. Levy, R. Shafiha, J. Fong, B.J. Luff and M. AsghariHigh, "Speed-GeSi-Electroabsorption Modulator on the SOI

- Waveguide Platform”, IEEE Journal of Selected Topics in Quantum Electronics, vol. 19, no. 6, 2013.
- [106] Filtre passe-bande optique, documentation disponible sur : <https://photonics.ixblue.com/files/files/pdf/CVF-200.pdf>, consultée le 23 novembre 2017.
- [107] Circulateur, documentation disponible sur : www.ozoptics.com/ALLNEW_PDF/DTS0070.pdf, consultée le 23 novembre 2017.
- [108] O. Chand, B. C. Choudhary and S. S. Pattnaik, “Characterization of Fiber-Bragg Grating as In-Line-Transmission Component in OFC Networks”, ICOP 2009-International Conference on Optics and Photonics CSIO, Chandigarh, India, 30 Oct.-1 Nov. 2009.
- [109] Duplexeur, documentation disponible sur : https://goochandhousego.com/wp-content/uploads/2017/01/GH_DS_FO_FFW_WDM_1310nm_or_1550nm_and_1625nm.pdf, consultée le 23 novembre 2017.
- [110] Multiplexeur et démultiplexeur de bandes C et L, documentation disponible sur : <http://www.sun-ins.com/lineup2/opneti/pdf-pm/wdm-filter/18\%20c-l\%20band\%20fdm.pdf>, consultée le 23 novembre 2017.
- [111] Multiplexeur et démultiplexeur de bandes C et L, documentation disponible sur : http://www.acphotonics.com/products/product_files/MWDM_CL-band.htm, consultée le 23 novembre 2017.
- [112] Recommandations ITU G983 et G984, disponibles sur : <https://www.itu.int>, consultées le 23 novembre 2017.
- [113] Wavelength Selective Switch Finisar, documentation disponible sur : <https://www.finisar.com/roadms-wavelength-management/10wsaaxxf11>, consultée le 23 novembre 2017.
- [114] K. Falta and C. Cameron, “Balancing Performance, Flexibility and Scalability in Optical Networks”, White paper, Finisar Corporation, 2012.
- [115] Wavelength Selective Switch JDSU, documentation disponible sur : <https://4donline.ihs.com/images/VipMasterIC/IC/JDSU/JDSUS00120/JDSUS00120-1.pdf?hkey=EF798316E3902B6ED9A73243A3159BB0>, consultée le 23 novembre 2017.
- [116] Multiplexeur flat-top, documentation disponible sur : <https://www.oquest.com/print/getproduct/18495/cat/1874/page>, consultée le 23 novembre 2017.
- [117] Amplificateur à fibre dopée à l’Erbium en bande de transmission C Keopsys, documentation disponible sur : <https://www.keopsys.com/portfolio/fiber-amplifier-c-pb-hp-series/>, consultée le 23 novembre 2017.
- [118] Amplificateur à fibre dopée à l’Erbium en bande de transmission L Keopsys, documentation disponible sur : <https://www.keopsys.com/portfolio/cw-erbium-fiber-amplifier-l-hg-series>, consultée le 23 novembre 2017.
- [119] C. R. Giles and E. Desurvire, “Modeling Erbium-Doped-Fiber Amplifiers”, Journal of Lightwave Technology, vol. 9., no. 2, pp. 271-283, February 1991.

- [120] A. K. Srivastava, J. L. Zyskind, Y. Sun, J. Ellson, G. Newsome, R. W. Tkach, A. R. Chraplyvy, J. W. Sulhoff, T. A. Strasser, C. Wolf, and J. R. Pedrazzani, "Fast-Link control Protection of Surviving Channels in Multiwavelength Optical Networks", *IEEE Photonics Technology Letters*, vol. 9, pp. 1667-1669, 1997.
- [121] S. R. Chinn, J. L. Zyskind, Y. Sun, J. Ellson, G. Newsome, R. W. Tkach, A. R. Chraplyvy, J. W. Sulhoff, T. A. Strasser, C. Wolf and J. R. Pedrazzani, "Simplified Modeling of Transients in Gain-Clamped Erbium-Doped-Fiber Amplifiers", *Journal of Lightwave Technology*, vol. 16, pp. 1095-1100, 1998.
- [122] C. Tian and S. Kinoshita, "Analysis and control of transient dynamics of EDFA pumped by 1480- and 980-nm lasers", *Journal of Lightwave Technology*, vol. 21, no. 8, pp. 1728-1734, 2003.
- [123] M. I. Hayee and A. E. Willner, "Fiber-Transmission Penalties Due to EDFA-Power-Transients Resulting from Fiber Nonlinearity and ASE Noise in Add/Drop-Multiplexed-WDM Networks", *Optical Fiber Communication Conference and the International Conference on Integrated Optics and Optical Fiber Communication OFC/IOOC '99*, 1999.
- [124] Optical Modulation Analyzer, documentation disponible sur : <https://www.dipaul.ru/upload/iblock/790/790f30ba6a30ca5cb960dd7a8152c923.pdf>, consultée le 23 novembre 2017.
- [125] Générateur de signaux pseudo-aléatoires, information disponible sur : <https://www.ecnmag.com/article/2011/06/raising-bar-advanced-signal-simulation>, consultée le 23 novembre 2017.
- [126] Iqtools, information disponible sur : <https://fr.mathworks.com/videos/agilent-m8190a-using-iqtools-and-matlab-91911.html>, consultée le 23 novembre 2017.
- [127] Filtre passe-bande optique reconfigurable, information disponible sur : <https://yenista.com/Tunable-Filters,8.html>, consultée le 23 novembre 2017.
- [128] R. A. Shafik, Md. S. Rahman, AHM R. Islam and N. S. Ashraf, "On the Error-Vector Magnitude as a Performance Metric and Comparative Analysis", *2nd International Conference on Emerging Technologies IEEE-ICET 2006*, Peshavar, Pakistan, 13-14 Nov. 2006.
- [129] M. Lin, Q. Zhang and Q. Xu, "EVM Simulation and its Comparison with BER for Different Types of Modulation", *TENCON 2007, 2007 IEEE Region 10 Conference*, Taipei, Taiwan, 30 Oct.-2 Nov. 2007.
- [130] W. Freude, R. Schmogrow, B. Nebedahl, M. Winter, A. Josten, D. Hillerkusss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker and J. Leuthold, "Quality Metrics for Optical Signals : Eye Diagram, Q-Factor, OSNR, EVM and BER", *14th International Conference on Transparent Optical Networks, ICTON 2012*, Coventry, England, 27-31 Jul. 2012.
- [131] B. Nebedahl, R. Schmogrow, T. Dennis, A. Josten, D. Hillerkusss, S. Koenig, J. Meyer, M. Dreschmann, M. Winter, M. Huebner, W. Freude, C. Koos and J. Leuthold, "Quality

- Metrics in Optical Modulation Analysis : EVM and its relation to Q-Factor, OSNR and BER”, 2012 Asia Communications and Photonics Conference, ACP 2012, Guangzhou, China, 10 Nov. 2012.
- [132] R. Schmogrow, B. Nebedahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, W. Freude, and J. Leuthold, “Error-Vector Magnitude as Performance Measure for Advanced Modulation Formats”, *IEEE Photonics Technology Letters*, vol. 24, no. 1, pp. 61-63, Jan. 2012.
- [133] M. Maxfield, “Gray Code Fundamentals - Part I”, *EE Times - Connecting the Global Electronics Community*, Internet blog, 28 May 2011, document disponible sur : https://www.eetimes.com/document.asp?doc_id=1278809, consulté le 23 novembre 2017.
- [134] S. Barthomeuf, F. Saliou, L. Anet Neto, P. Chanclou et D. Erasme, « Simplification du codage et décodage PAM4 Gray pour les réseaux d'accès optique au-delà de 10 Gbit/s », Journées nationales d'optique guidée, JNOG 2017, Limoges, France, 3-6 Jul. 2017.
- [135] J. Dong, “Estimation of Bit Error Rate of any digital Communication System”, thèse pour obtenir le grade de docteur délivré par Télécom Bretagne, 18 Déc. 2013.
- [136] “Soft-Decision FEC Benefits for 100G”, Fujitsu document, 2012. Document disponible sur : <http://www.fujitsu.com/downloads/TEL/fnc/whitepapers/Soft-Decision-FEC-Benefits-or-100G-wp.pdf>, consulté le 23 novembre 2017.
- [137] “Soft-decision Forward Error Correction for Coherent Super-channels”, white paper Infinera, Document Number : WP-SD-FEC-4-2016, 2016. Document disponible sur : https://www.infinera.com/wp-content/uploads/2015/08/WP-SD-FEC_for_Coherent_Super-Channels.pdf, consulté le 23 novembre 2017.
- [138] H. Kawanishi, Y. Yamauchi, N. Mineo, Y. Shibuya, H. Murai, K. Yamada and H. Wada, “EAM-Integrated DFB Laser Modules with More Than 40-GHz Bandwidth”, *IEEE Photonics Technology Letters*, vol. 13, no. 9, pp. 954-2219, Sep. 2001.
- [139] S. Dris, J. Bauwelinck, X. Yin, B. Schrenk, J.A. Lazaro, V. Katopodis, P. Bakopoulos, J. Verbrugghe and H. Avramopoulos, “Quaternary TDM-PAM and its Implications for TDMA Equipment”, 2012 38th European Conference and Exhibition on Optical Communications, Amsterdam, Netherlands, 16-20 Sept. 2012.
- [140] M. Chagnon, M. Morsy-Osman, M. Poulin, C. Paquet, S. Lessard and D. V. Plant, “Experimental-Parametric Study of a Silicon-Photonic Modulator enabled 112-Gb/s-PAM-Transmission System with a DAC and ADC”, *Journal of Lightwave Technology*, vol. 33, no. 7, pp. 1380-1387, Apr. 2015.

GLOSSAIRE

- AM : Amplitude Modulation
- APD : Avalanche Photodiode
- ASE : Amplified-Spontaneous Emission
- ASK : Amplitude-Shift Keying
- AWG : Arrayed-Waveguide Grating
- BER : Bit-Error Rate
- BPF : Band-Pass Filter
- BVT : Bandwidth-Variable Transceiver
- DC : Data Center
- DBR : Distributed Bragg Reflector
- DFB : Distributed Feedback
- DOS : Data-Center-Optical Switch
- DSP : Digital-Signal Processing
- EAM : Electro-Absorption Modulator
- EODCN : Elastic-Optical-Data-Center Network
- EPOC : Energy-Proportional and Opportunistic Computing
- E-POPI : Extended-Passive-Optical-Pod Interconnect
- EPS : Electrical-Packet Switching
- EDFA : Erbium-Doped-Fiber Amplifier
- EVM : Error-Vector Magnitude
- FDL : Fiber-Delay Line
- FEC : Forward-Error Correction
- FBG : Fiber-Bragg Grating

- FP : Fabry-Perot
- FWM : Four-Wave Mixing
- GES : Gaz à effet serre
- GeSi : The Climate Group on behalf of the Global eSustainability Initiative
- HD : Header Detector
- HD-FEC : Hard-Decision-Forward-Error Correction
- HELIOS : Hybrid-Electrical/Optical Switch
- IP : Internet Protocol
- IRIS : Integrated-Router Interconnected Spectrally
- ITU : International-Telecommunications Union
- MEMS : Micro-Electro-mechanical-Mirrors Systems
- MMF : Multi-Mode Fiber
- MZM : Mach-Zehnder Modulator
- NRZ : Non Return to Zero
- OBS : Optical-Burst Switching
- OCS : Optical-Circuit Switching
- O-E-O : Optical-Electrical-Optical
- OI : Optical Interface
- OFDM : Orthogonal-Frequency-Division Multiplexing
- OMA : Optical Modulation Analyzer
- ONI : Optical-Network Interface
- OOK : On-Off Keying
- OPS : Optical-Packet Switching
- OSA : Optical Spectrum Analyzer
- OSNR : Optical-Signal-to-Noise Ratio
- OSS : Optical Selective Switch
- PAM : Pulse-Amplitude Modulation
- PDM : Polarization-Division Multiplexing
- PMD : Polarization-Mode Dispersion
- POADM : Packet-Optical-Add and Drop Multiplexer

-
- POITR : Passive-Optical Interconnect at Top of the Rack
 - PON : Passive-Optical Networks
 - POPI : Passive-Optical-Pod Interconnect
 - PSK : Phase-Shift Keying
 - QAM : Quadrature-Amplitude Modulation
 - RZ : Return to Zero
 - RIN : Relative-Intensity Noise
 - SAN : Storage Area Network
 - SDN : Software-Defined Networking
 - SER : Symbol-Error Rate
 - SD-FEC : Soft-Decision-Forward-Error Correction
 - SG : Sampled Grating
 - SMF : Single-Mode Fiber
 - SOA : Semiconductor-Optical Amplifier
 - SPM : Self-Phase Modulation
 - SSG : Super-Structure Grating
 - SSMF : Standard-Single-Mode Fiber
 - TB : Time Buffer
 - TCAM : Ternary-Content-Addressable Memory
 - T-POPI : Transfert-Passive-Optical-Pod Interconnect
 - ToR : Top of the Rack
 - TWDM : Time-and Wavelength-Division Multiplexing
 - TWDM-PON : Time-and Wavelength-Division Multiplexing - Passive-Optical Networks
 - TWIN : Time-domain-Wavelength-Interleaved Networks
 - VM : Virtual Machine
 - VCSEL : Vertical-Cavity-Surface-Emitting Laser
 - WC : Wavelength Converter
 - WDM : Wavelength-Division Multiplexing
 - WS : Wavelength Switch
 - WSS : Wavelength-Selective Switch

— XPM : Cross-Phase Modulation

DIMENSIONNEMENT DES RÉSEAUX POPI ET E-POPI

Le nombre maximal \hat{N} de serveurs connectés au réseau POPI est calculé pour les formats de modulation PAM 4 et 8 à des débits de 14, 28, 56 et 112 Gbit/s. Ces résultats sont présentés dans les tableaux B.1, B.2, B.3 et B.4, respectivement. Ainsi que le nombre maximal \hat{N} de serveurs connectés au réseau E-POPI, pour les mêmes formats de modulation et débits est présenté dans les tableaux B.5, B.6, B.7 et B.8.

Nos calculs considèrent différentes tailles de grille (de 25 à 100 GHz avec un pas de 12,5 GHz) selon la bande passante du signal et la faisabilité du filtre optique passe-bande nécessaire. Certains résultats sont présentés en couleur bleue ou rouge, cette couleur représente la contrainte la plus importante. Bleue : plage d'accord et largeur de canal ($\hat{N}_{X-GHx} < \hat{N}_{IL}$) ; et rouge : pertes d'insertion ($\hat{N}_{IL} < \hat{N}_{X-GHx}$). Tandis que les valeurs présentées en couleur violette représentent le fait que ces deux limitations soient égales ($\hat{N}_{IL} = \hat{N}_{X-GHx}$). Dans l'idéal, les deux limitations devraient être le plus proches possible ($\hat{N}_{IL} \approx \hat{N}_{X-GHz}$), voir égales, afin de maximiser le nombre de serveurs connectés tout en utilisant une puissance en émission minimale.

Dimensionnement de POPI à 14 Gbit/s						
Format de modulation	PAM 4			PAM 8		
Récepteur	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
P_{Rx} (dBm)	-18,6	-25,4	-28,1	-13,8	-18,6	-24,5
P_{Tx} (dBm)	15					
\hat{L}_C (dB)	18,1	23,9	26,2	14	18,1	23,1
\hat{N}_{IL}	64	246	419	24	64	206
\hat{N} avec $X = 25$ GHz	64	125	125	24	64	125
\hat{N} avec $X = 37,5$ GHz	64	82	82	24	64	82
\hat{N} avec $X = 50$ GHz	61	61	61	24	61	61
\hat{N} avec $X = 62,5$ GHz	48	48	48	24	48	48
\hat{N} avec $X = 75$ GHz	39	39	39	24	39	39
\hat{N} avec $X = 87,5$ GHz	33	33	33	24	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	24	29	29
P_{Tx} (dBm)	20					
\hat{L}_C (dB)	22,4	28,2	30,5	18,3	22,4	27,4
\hat{N}_{IL}	172	660	1125	66	172	553
\hat{N} avec $X = 25$ GHz	125	125	125	66	125	125
\hat{N} avec $X = 37,5$ GHz	82	82	82	66	82	82
\hat{N} avec $X = 50$ GHz	61	61	61	61	61	61
\hat{N} avec $X = 62,5$ GHz	48	48	48	48	48	48
\hat{N} avec $X = 75$ GHz	39	39	39	39	39	39
\hat{N} avec $X = 87,5$ GHz	33	33	33	33	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	29	29	29
P_{Tx} (dBm)	25					
\hat{L}_C (dB)	26,7	32,5	34,8	22,5	26,7	31,7
\hat{N}_{IL}	463	1772	3019	179	463	1483
\hat{N} avec $X = 25$ GHz	125	125	125	125	125	125
\hat{N} avec $X = 37,5$ GHz	82	82	82	82	82	82
\hat{N} avec $X = 50$ GHz	61	61	61	61	61	61
\hat{N} avec $X = 62,5$ GHz	48	48	48	48	48	48
\hat{N} avec $X = 75$ GHz	39	39	39	39	39	39
\hat{N} avec $X = 87,5$ GHz	33	33	33	33	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	29	29	29

Tableau B.1. Nombre maximal de serveurs connectés au réseau POPI pour les formats de modulation PAM 4 et 8 à 14 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $BER = 10^{-3}$.

Dimensionnement de POPI à 28 Gbit/s						
Format de modulation	PAM 4			PAM 8		
Récepteur	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
P_{Rx} (dBm)	-16,6	-24,5	-25	-12,1	-16,3	-20,7
P_{Tx} (dBm)	15					
\hat{L}_C (dB)	16,4	22,9	23,6	12,5	16,1	19,9
\hat{N}_{IL}	43	194	227	17	40	97
\hat{N} avec $X = 25$ GHz	NA	NA	NA	17	40	97
\hat{N} avec $X = 37,5$ GHz	43	82	82	17	40	82
\hat{N} avec $X = 50$ GHz	43	61	61	17	40	61
\hat{N} avec $X = 62,5$ GHz	43	48	48	17	40	48
\hat{N} avec $X = 75$ GHz	39	39	39	17	39	39
\hat{N} avec $X = 87,5$ GHz	33	33	33	17	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	17	29	29
P_{Tx} (dBm)	20					
\hat{L}_C (dB)	20,7	27,2	27,9	16,8	20,4	24,2
\hat{N}_{IL}	116	521	610	47	109	261
\hat{N} avec $X = 25$ GHz	NA	NA	NA	47	109	125
\hat{N} avec $X = 37,5$ GHz	82	82	82	47	82	82
\hat{N} avec $X = 50$ GHz	61	61	61	47	61	61
\hat{N} avec $X = 62,5$ GHz	48	48	48	47	48	48
\hat{N} avec $X = 75$ GHz	39	39	39	39	39	39
\hat{N} avec $X = 87,5$ GHz	33	33	33	33	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	29	29	29
P_{Tx} (dBm)	25					
\hat{L}_C (dB)	24,9	31,5	32,1	21,1	24,7	28,5
\hat{N}_{IL}	312	1398	1637	128	294	700
\hat{N} avec $X = 25$ GHz	NA	NA	NA	125	125	125
\hat{N} avec $X = 37,5$ GHz	82	82	82	82	82	82
\hat{N} avec $X = 50$ GHz	61	61	61	61	61	61
\hat{N} avec $X = 62,5$ GHz	48	48	48	48	48	48
\hat{N} avec $X = 75$ GHz	39	39	39	39	39	39
\hat{N} avec $X = 87,5$ GHz	33	33	33	33	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	29	29	29

NA : N'est pas applicable. Représente les cas où le filtre passe-bande requis n'est pas réalisable, car il exige une pente supérieure à 800 dB/nm.

Tableau B.2. Nombre maximal de serveurs connectés au réseau POPI pour les formats de modulation PAM 4 et 8 à 28 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $BER = 10^{-3}$.

Dimensionnement de POPI à 56 Gbit/s						
Format de modulation	PAM 4			PAM 8		
Récepteur	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
P_{Rx} (dBm)	-14,3	-21,9	-24,8	-9,4	-14,1	-17,2
P_{Tx} (dBm)	15					
\hat{L}_C (dB)	14,4	20,9	23,4	10,2	14,2	16,9
\hat{N}_{IL}	27	123	218	10	26	48
\hat{N} avec $X = 37,5$ GHz	NA	NA	NA	10	26	48
\hat{N} avec $X = 50$ GHz	27	61	61	10	26	48
\hat{N} avec $X = 62,5$ GHz	27	48	48	10	26	48
\hat{N} avec $X = 75$ GHz	27	39	39	10	26	39
\hat{N} avec $X = 87,5$ GHz	27	33	33	10	26	33
\hat{N} avec $X = 100$ GHz	27	29	29	10	26	29
P_{Tx} (dBm)	20					
\hat{L}_C (dB)	18,6	25,2	27,7	14,5	18,5	21,2
\hat{N}_{IL}	73	331	586	28	71	130
\hat{N} avec $X = 37,5$ GHz	NA	NA	NA	28	71	82
\hat{N} avec $X = 50$ GHz	61	61	61	28	61	61
\hat{N} avec $X = 62,5$ GHz	48	48	48	28	48	48
\hat{N} avec $X = 75$ GHz	39	39	39	28	39	39
\hat{N} avec $X = 87,5$ GHz	33	33	33	28	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	28	29	29
P_{Tx} (dBm)	25					
\hat{L}_C (dB)	23	29,5	32	18,8	22,8	25,5
\hat{N}_{IL}	198	888	1574	75	190	351
\hat{N} avec $X = 37,5$ GHz	NA	NA	NA	75	82	82
\hat{N} avec $X = 50$ GHz	61	61	61	61	61	61
\hat{N} avec $X = 62,5$ GHz	48	48	48	48	48	48
\hat{N} avec $X = 75$ GHz	39	39	39	39	39	39
\hat{N} avec $X = 87,5$ GHz	33	33	33	33	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	29	29	29

NA : N'est pas applicable. Représente les cas où le filtre passe-bande requis n'est pas réalisable, car il exige une pente supérieure à 800 dB/nm.

Tableau B.3. Nombre maximal de serveurs connectés au réseau POPI pour les formats de modulation PAM 4 et 8 à 56 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $BER = 10^{-3}$.

Dimensionnement de POPI à 112 Gbit/s						
Format de modulation	PAM 4			PAM 8		
Récepteur	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
P_{Rx} (dBm)	-10,3	-16	-21,7	-5,9	-8,7	-13
P_{Tx} (dBm)	15					
\hat{L}_C (dB)	11	15,9	20,7	7,2	9,6	13,3
\hat{N}_{IL}	12	38	118	5	9	21
\hat{N} avec $X = 62,5$ GHz	NA	NA	NA	5	9	21
\hat{N} avec $X = 75$ GHz	NA	NA	NA	5	9	21
\hat{N} avec $X = 87,5$ GHz	NA	NA	NA	5	9	21
\hat{N} avec $X = 100$ GHz	12	29	29	5	9	21
P_{Tx} (dBm)	20					
\hat{L}_C (dB)	15,3	20,1	25	11,5	13,9	17,6
\hat{N}_{IL}	33	103	318	14	24	57
\hat{N} avec $X = 62,5$ GHz	NA	NA	NA	14	24	48
\hat{N} avec $X = 75$ GHz	NA	NA	NA	14	24	39
\hat{N} avec $X = 87,5$ GHz	NA	NA	NA	14	24	33
\hat{N} avec $X = 100$ GHz	29	29	29	14	24	29
P_{Tx} (dBm)	25					
\hat{L}_C (dB)	19,5	24,4	29,3	15,8	18,2	21,9
\hat{N}_{IL}	90	277	853	37	65	153
\hat{N} avec $X = 62,5$ GHz	NA	NA	NA	37	48	48
\hat{N} avec $X = 75$ GHz	NA	NA	NA	37	39	39
\hat{N} avec $X = 87,5$ GHz	NA	NA	NA	33	33	33
\hat{N} avec $X = 100$ GHz	29	29	29	29	29	29

NA : N'est pas applicable. Représente les cas où le filtre passe-bande requis n'est pas réalisable, car il exige une pente supérieure à 800 dB/nm.

Tableau B.4. Nombre maximal de serveurs connectés au réseau POPI pour les formats de modulation PAM 4 et 8 à 112 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à BER = 10^{-3} .

Dimensionnement d'E-POPI à 14 Gbit/s						
Format de modulation	PAM 4			PAM 8		
Récepteur	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
P_{Rx} (dBm)	-18,6	-25,4	-28,1	-13,8	-18,6	-24,5
P_{Tx} (dBm)	15					
\hat{L}_C (dB)	17,2	23,1	25,4	13,1	17,2	22,3
\hat{N}_{IL}	52	202	344	20	52	169
\hat{N} avec $X = 25$ GHz	52	202	253	20	52	169
\hat{N} avec $X = 37,5$ GHz	52	167	167	20	52	167
\hat{N} avec $X = 50$ GHz	52	125	125	20	52	125
\hat{N} avec $X = 62,5$ GHz	52	99	99	20	52	99
\hat{N} avec $X = 75$ GHz	52	82	82	20	52	82
\hat{N} avec $X = 87,5$ GHz	52	70	70	20	52	70
\hat{N} avec $X = 100$ GHz	52	61	61	20	52	61
P_{Tx} (dBm)	20					
\hat{L}_C (dB)	21,5	27,3	29,7	17,4	21,5	26,6
\hat{N}_{IL}	141	542	924	54	141	454
\hat{N} avec $X = 25$ GHz	141	253	253	54	141	253
\hat{N} avec $X = 37,5$ GHz	141	167	167	54	141	167
\hat{N} avec $X = 50$ GHz	125	125	125	54	125	125
\hat{N} avec $X = 62,5$ GHz	99	99	99	54	99	99
\hat{N} avec $X = 75$ GHz	82	82	82	54	82	82
\hat{N} avec $X = 87,5$ GHz	70	70	70	54	70	70
\hat{N} avec $X = 100$ GHz	61	61	61	54	61	61
P_{Tx} (dBm)	25					
\hat{L}_C (dB)	25,8	31,6	33,9	21,7	25,8	30,9
\hat{N}_{IL}	380	1454	2479	147	380	1218
\hat{N} avec $X = 25$ GHz	253	253	253	147	253	253
\hat{N} avec $X = 37,5$ GHz	167	167	167	147	167	167
\hat{N} avec $X = 50$ GHz	125	125	125	125	125	125
\hat{N} avec $X = 62,5$ GHz	99	99	99	99	99	99
\hat{N} avec $X = 75$ GHz	82	82	82	82	82	82
\hat{N} avec $X = 87,5$ GHz	70	70	70	70	70	70
\hat{N} avec $X = 100$ GHz	61	61	61	61	61	61

Tableau B.5. Nombre maximal de serveurs connectés au réseau E-POPI pour les formats de modulation PAM 4 et 8 à 14 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $BER = 10^{-3}$.

Dimensionnement d'E-POPI à 28 Gbit/s						
Format de modulation	PAM 4			PAM 8		
Récepteur	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
P_{Rx} (dBm)	-16,6	-24,5	-25	-12,1	-16,3	-20,7
P_{Tx} (dBm)	15					
\hat{L}_C (dB)	15,5	22	22,7	11,7	15,3	19
\hat{N}_{IL}	35	159	186	14	33	79
\hat{N} avec $X = 25$ GHz	NA	NA	NA	14	33	79
\hat{N} avec $X = 37,5$ GHz	35	159	167	14	33	79
\hat{N} avec $X = 50$ GHz	35	125	125	14	33	79
\hat{N} avec $X = 62,5$ GHz	35	99	99	14	33	79
\hat{N} avec $X = 75$ GHz	35	82	82	14	33	79
\hat{N} avec $X = 87,5$ GHz	35	70	70	14	33	70
\hat{N} avec $X = 100$ GHz	35	61	61	14	33	61
P_{Tx} (dBm)	20					
\hat{L}_C (dB)	19,8	26,3	27	15,9	19,5	23,3
\hat{N}_{IL}	95	427	501	39	90	214
\hat{N} avec $X = 25$ GHz	NA	NA	NA	39	90	214
\hat{N} avec $X = 37,5$ GHz	95	167	167	39	90	167
\hat{N} avec $X = 50$ GHz	95	125	125	39	90	125
\hat{N} avec $X = 62,5$ GHz	95	99	99	39	90	99
\hat{N} avec $X = 75$ GHz	82	82	82	39	82	82
\hat{N} avec $X = 87,5$ GHz	70	70	70	39	70	70
\hat{N} avec $X = 100$ GHz	61	61	61	39	61	61
P_{Tx} (dBm)	25					
\hat{L}_C (dB)	24,1	30,6	31,3	20,2	23,8	27,6
\hat{N}_{IL}	256	1148	1344	105	241	575
\hat{N} avec $X = 25$ GHz	NA	NA	NA	105	241	253
\hat{N} avec $X = 37,5$ GHz	167	167	167	105	167	167
\hat{N} avec $X = 50$ GHz	125	125	125	105	125	125
\hat{N} avec $X = 62,5$ GHz	99	99	99	99	99	99
\hat{N} avec $X = 75$ GHz	82	82	82	82	82	82
N avec $X = 87,5$ GHz	70	70	70	70	70	70
N avec $X = 100$ GHz	61	61	61	61	61	61

NA : N'est pas applicable. Représente les cas où le filtre passe-bande requis n'est pas réalisable, car il exige une pente supérieure à 800 dB/nm.

Tableau B.6. Nombre maximal de serveurs connectés au réseau E-POPI pour les formats de modulation PAM 4 et 8 à 28 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $BER = 10^{-3}$.

Dimensionnement d'E-POPI à 56 Gbit/s						
Format de modulation	PAM 4			PAM 8		
Récepteur	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
P_{Rx} (dBm)	-14,3	-21,9	-24,8	-9,4	-14,1	-17,2
P_{Tx} (dBm)	15					
\hat{L}_C (dB)	13,5	20,1	22,5	9,3	13,4	16
\hat{N}_{IL}	22	101	179	8	21	40
\hat{N} avec $X = 37,5$ GHz	NA	NA	NA	8	21	40
\hat{N} avec $X = 50$ GHz	22	101	125	8	21	40
\hat{N} avec $X = 62,5$ GHz	22	99	99	8	21	40
\hat{N} avec $X = 75$ GHz	22	82	82	8	21	40
\hat{N} avec $X = 87,5$ GHz	22	70	70	8	21	40
\hat{N} avec $X = 100$ GHz	22	61	61	8	21	40
P_{Tx} (dBm)	20					
\hat{L}_C (dB)	17,8	24,3	26,8	13,6	17,7	20,3
\hat{N}_{IL}	60	271	481	23	58	107
\hat{N} avec $X = 37,5$ GHz	NA	NA	NA	23	58	107
\hat{N} avec $X = 50$ GHz	60	125	125	23	58	107
\hat{N} avec $X = 62,5$ GHz	60	99	99	23	58	99
\hat{N} avec $X = 75$ GHz	60	82	82	23	58	82
\hat{N} avec $X = 87,5$ GHz	60	70	70	23	58	70
\hat{N} avec $X = 100$ GHz	60	61	61	23	58	61
P_{Tx} (dBm)	25					
\hat{L}_C (dB)	22,1	28,6	31,1	17,9	21,9	24,6
\hat{N}_{IL}	162	729	1292	61	156	288
\hat{N} avec $X = 37,5$ GHz	NA	NA	NA	61	167	167
\hat{N} avec $X = 50$ GHz	125	125	125	61	125	125
\hat{N} avec $X = 62,5$ GHz	99	99	99	61	99	99
\hat{N} avec $X = 75$ GHz	82	82	82	61	82	82
\hat{N} avec $X = 87,5$ GHz	70	70	70	61	70	70
\hat{N} avec $X = 100$ GHz	61	61	61	61	61	61

NA : N'est pas applicable. Représente les cas où le filtre passe-bande requis n'est pas réalisable, car il exige une pente supérieure à 800 dB/nm.

Tableau B.7. Nombre maximal de serveurs connectés au réseau E-POPI pour les formats de modulation PAM 4 et 8 à 56 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à $BER = 10^{-3}$.

Dimensionnement d'E-POPI à 112 Gbit/s						
Format de modulation	PAM 4			PAM 8		
Récepteur	PIN	APD	SOA-PIN	PIN	APD	SOA-PIN
P_{Rx} (dBm)	-10,3	-16	-21,7	-5,9	-8,7	-13
P_{Tx} (dBm)	15					
\hat{L}_C (dB)	10,1	15	19,9	6,3	8,7	12,4
\hat{N}_{IL}	10	31	97	4	7	17
\hat{N} avec $X = 62,5$ GHz	NA	NA	NA	4	7	17
\hat{N} avec $X = 75$ GHz	NA	NA	NA	4	7	17
\hat{N} avec $X = 87,5$ GHz	NA	NA	NA	4	7	17
\hat{N} avec $X = 100$ GHz	10	31	61	4	7	17
P_{Tx} (dBm)	20					
\hat{L}_C (dB)	14,4	19,3	24,2	10,6	13	16,7
\hat{N}_{IL}	27	84	261	11	20	46
\hat{N} avec $X = 62,5$ GHz	NA	NA	NA	11	20	46
\hat{N} avec $X = 75$ GHz	NA	NA	NA	11	20	46
\hat{N} avec $X = 87,5$ GHz	NA	NA	NA	11	20	46
\hat{N} avec $X = 100$ GHz	27	61	61	11	20	46
P_{Tx} (dBm)	25					
\hat{L}_C (dB)	18,7	23,6	28,5	14,9	17,3	21
\hat{N}_{IL}	73	227	700	31	53	125
\hat{N} avec $X = 62,5$ GHz	NA	NA	NA	31	53	99
\hat{N} avec $X = 75$ GHz	NA	NA	NA	31	53	82
\hat{N} avec $X = 87,5$ GHz	NA	NA	NA	31	53	70
\hat{N} avec $X = 100$ GHz	61	61	61	31	53	61

NA : N'est pas applicable. Représente les cas où le filtre passe-bande requis n'est pas réalisable, car il exige une pente supérieure à 800 dB/nm.

Tableau B.8. Nombre maximal de serveurs connectés au réseau E-POPI pour les formats de modulation PAM 4 et 8 à 112 Gbit/s, pour les récepteurs PIN, APD et SOA-PIN, à des puissances d'émission de 12, 20 et 25 dBm, pour une limite du FEC à BER = 10^{-3} .

PUBLICATIONS

- B. Dumas Feris, P. Gravey, P. Morel, M-L. Moulinard, M. Morvan and A. Sharaiha, “Dimensioning of 112G Optical-Packet-Switching-Based Interconnects for Energy-Efficient Data Centers”, Optical-Network Design and Modeling (ONDM) special issue of Journal of Optical Communications and Networking (JOCN), vol. 9, no. 4, pp. B124-B136, avril 2017.
- B. Dumas Feris, P. Gravey, P. Morel, M-L. Moulinard, M. Morvan and A. Sharaiha, “Pulse Amplitude Modulation Applied to Extended Passive Optical Pod Interconnect, for Small Energy-Aware Data Centers”, présenté à ICTON 2016 (18th International Conference on Transparent Optical Networks), 10-14 juillet 2016, Trento, Italie.
- B. Dumas Feris, P. Gravey, P. Morel, M-L. Moulinard, M. Morvan and A. Sharaiha, “112-Gbit/s Passive Optical Pod Interconnect for Small Data Centers Using Pulse Amplitude Modulation”, présenté à ONDM 2016 (20th International Conference on Optical Network Design and Modeling), mai 2016, Cartagena, Espagne.
- N. Beldiceanu, B. Dumas Feris, P. Gravey, S. Hasan, C. Jard, T. Ledoux, Y. Li, D. Lime, G. Madi-Wamba, J-M. Menaud, P. Morel, M. Morvan, M-L. Moulinard, A-C. Orgerie, J-L. Pazat, O. Roux and A. Sharaiha, “Towards Energy-Proportional Clouds Partially Powered by Renewable Energy”, Special Issue, Journal Computing (Springer), décembre 2015.
- B. Dumas Feris, P. Morel, J. Pesic, P. Gravey, M-L. Moulinard, M. Morvan and A. Sharaiha, “Extended Passive Optical Pod Interconnect for Small Mono-site Data Centers”, présenté à ICTON 2015 (17th International Conference on Transparent Optical Networks), 5-9 juillet 2015, Budapest, Hongrie.
- N. Beldiceanu, B. Dumas Feris, P. Gravey, S. Hasan, C. Jard, T. Ledoux, Y. Li, D. Lime, G. Madi-Wamba, J-M. Menaud, P. Morel, M. Morvan, M-L. Moulinard, A-C. Orgerie, J-L. Pazat, O. Roux and A. Sharaiha, “The EPOC Project : Energy Proportional and Opportunistic Computing System”, Position paper, présenté à SMART-GREENS 2015 (4th International Conference on Smart Cities and Green ICT systems), 20-22 mai 2015, Lisbonne, Portugal.

La consommation d'énergie des centres de données est un enjeu majeur. Leurs communications internes représentent près du quart de cette consommation. Les technologies de commutation optique ont en principe une meilleure efficacité énergétique que les solutions actuelles. Ce travail porte sur les réseaux optiques en mode paquet pour des centres de données de petite et moyenne taille. Il s'est déroulé dans le cadre du projet EPOC (*Energy Proportional and Opportunistic Computing*) qui s'intéresse à la réduction de la consommation d'énergie d'un centre de données alimenté partiellement par des sources renouvelables. Une hypothèse clé est l'absence d'un réseau de stockage dédié réduisant ainsi la consommation des interconnexions. Par contre, afin de pouvoir éteindre certains serveurs selon la charge de travail et l'énergie disponible, le débit doit être de l'ordre de 100 Gbit/s. Après un état de l'art des réseaux optiques pour centre de données nous choisissons une solution reposant sur une infrastructure entièrement passive et des émetteurs rapidement accordables en longueur d'onde, proposée récemment dans la littérature (POPI).

Nous étudions les limitations dues aux composants (pertes d'insertion, plage d'accord et espacement des canaux). Nous proposons une extension (E-POPI) qui permet d'augmenter le nombre de serveurs connectés en utilisant plusieurs plages de longueurs d'onde. Pour les centres de données de plus grande taille nous proposons un réseau à deux étages (intra- et inter-racks) opérant respectivement dans les bandes C et L, POPI+. La connexion entre étages se fait via une passerelle transparente dans un sens et opaque dans l'autre. Différentes solutions de contrôle des deux étages sont détaillées.

Un des éléments essentiels de faisabilité de ces architectures est de concilier la montée en débit avec les pertes du réseau passif d'interconnexion. Les techniques cohérentes des transmissions longue distance ne sont pas actuellement envisageables pour un centre de données. Nous avons donc étudié les formats PAM 4 et 8, par simulation avec différents débits (jusqu'à 112 Gbit/s et récepteurs (PIN, APD et SOA-PIN) et aussi, expérimentalement, à 12 et 18 Gbit/s. Nous avons développé une méthode de compensation des distorsions générées par les différents composants qui procure un compromis entre précision de correction et temps de calcul.

Ces résultats nous permettent de déterminer les pertes d'insertion tolérables. Nous les combinons avec les limitations liées à la plage d'accord des émetteurs et à l'encombrement spectral des canaux occupant des fenêtres multiples de 12,5 GHz pour dimensionner les différentes architectures. Les réseaux POPI, E-POPI et POPI+ permettent respectivement la connexion de 48, 99 et 2352 entités à 112 Gbit/s. Nos évaluations tiennent compte d'une possible dispersion des caractéristiques des principaux composants.

Mots clés : Réseaux optiques, Centre de données, Commutation de paquets optiques, Formats de modulation d'amplitude PAM 4 et PAM 8

Data-center energy consumption is nowadays a major issue. Intra-data-center networking accounts almost for a quarter of the data-center total power consumption. Optical switching technologies could provide higher power efficiency than current solutions based on electrical-packet switching. This work focuses on optical-packet-switched networks for small- and medium-size data centers. It takes part of the EPOC (*Energy-Proportional and Opportunistic Computing*) project, which main interest consists on reducing the overall power consumption of a data center partially powered by renewable sources. A key assumption is that our data center does not rely on a dedicated storage network, in order to reduce the consumption of those interconnections. In addition, with the aim of being able to turn off some servers according to the workload and the available energy, the bit rate must be close to 100 Gbit/s. We have chosen, after studying the state of the art of data-center interconnects, a purely passive network architecture based on fast-wavelength-tunable transmitters under the name of POPI.

We study POPI's limitations due to its components (insertion loss, tuning range and channel spacing). We then propose an extension called E-POPI that allows to increase the number of connected servers by using several transmission bands. For larger data centers, we propose POPI+, a two-stage infrastructure for intra- and inter-rack communications operating in the C and L bands, respectively. The connection between both stages is done via a transparent gateway in one direction and an opaque one in the other. We discuss different control solutions for both stages.

The feasibility of these architectures depends on, among other factors, dealing with bit-rate increasing and power losses of a passive interconnect. Coherent long-distance-transmission techniques are not currently suited to data centers. We therefore studied PAM 4 and 8 modulation formats with direct detection. On one hand, by simulation, with different bit rates (up to 112 Gbit/s) and receivers (PIN, APD and SOA-PIN) and, on the other hand, experimentally, at 12 and 18 Gbit/s. We have developed a method for compensating the distortions generated by the different network components. Our method takes into account a good tradeoff between correction accuracy and computation time.

Simulation results allow us to determine the amount of insertion loss that may be supported. We then combine these results with the limitations of transmitters-tuning range and channel spacing using multiple of 12.5 GHz slots for dimensioning the proposed architectures. POPI, E-POPI and POPI+ interconnects allow the connection of 48, 99 and 2352 entities, respectively, at 112 Gbit/s. Our assessments take into account a potential dispersion of the characteristics of the main architecture components.

Keywords: Optical network, Data center, Optical-packet switching, Pulse-amplitude modulation PAM 4 and PAM 8

N° d'ordre : 2017IMTA0057

IMT Atlantique Bretagne-Pays de la Loire - www.imt-atlantique.fr

Campus de Brest
Technopôle Brest-Iroise
CS 83818
29238 Brest Cedex 03
T +33 (0)2 29 00 11 11
F +33 (0)2 29 00 10 00

Campus de Nantes
4, rue Alfred Kastler - La Chantrerie
CS 20722
44307 Nantes Cedex 3
T +33 (0)2 51 85 81 00
F +33 (0)2 51 85 81 99

Campus de Rennes
2, rue de la Châtaigneraie
CS 17607
35576 Cesson Sévigné Cedex
T +33 (0)2 99 12 70 00
F +33 (0)2 99 12 70 08