Garantie de la qualité de service et évaluation de performance des réseaux de télécommunications

Joanna Tomasik

Université Paris Sud
UFR scientifique d'Orsay

Garantie de la qualité de service et évaluation de performance des réseaux de télécommunications

mémoire pour
l'habilitation à diriger les recherches
présenté par

Joanna TOMASIK

Supélec Sciences des Systèmes (E3S)
Département Informatique

soutenue le 4 janvier 2012 devant le jury :

Khaldoun Al Agha	Université Paris-Sud 11
Tadeusz Czachórski	Institut de l'Informatique Théorique et Appliquée de l'Académie des Sciences de Pologne, Gliwice, Pologne
Andrzej Duda	INP-Ensimag (rapporteur)
Andrzej Jajszczyk	Université des Sciences et Technologie, Cracovie, Pologne (rapporteur)
Olivier Marcé	Alcatel-Lucent Bell Labs France (invité)
Sébastien Tixeuil	Université Pierre et Marie Curie, Paris 6 (rapporteur)
Guy Vidal-Naquet	Supélec et Université Paris-Sud 11

Table des matières

Table des symboles

Voici l'ensemble des symboles, notations, abbreviations et des conventions utilisés.

$\mathbb{R}$	L'ensemble des nombres réels
$\mathbb{R}^+$	L'ensemble des nombres réels positifs
$\mathbb{R}_+^*$	L'ensemble des nombres réels strictement positifs
$\mathbb{N}$	L'ensemble des nombres entiers naturels
$\mathbb{N}^*$	L'ensemble des nombres entiers naturels strictement positifs
$G = (V, E)$	Un graphe non-orienté avec V et E, les ensembles des nœuds et des arêtes
$G = (V, A)$	Un graphe orient avec V et A les ensembles des nœuds et E des arcs
$G = (V, A, E)$	Un graphe mixte avec V, A et E les ensembles des nœuds, des arcs et des arêtes
$[a, b]$	Une arête entre a et b
(a, b)	Un arc de a vers b
$< x, y, z >$	Une route x, y, z
P2P	Une relation de pair à pair, *peer-to-peer*
C2P	Une relation de client à fournisseur, *customer-to-provider*

DAG	*Directed Acyclic Graph*, un graphe orienté sans ciruit
GSMP	*Generalised Semi-Markov Process*
MMPP	*Markov Modulated Poisson Process*, processus de Poisson modulé par une chaîne de Markov
PEPA	*Performance Evaluation Process Algebra*
SAN, RAS	*Stochastic Automata Network*, Réseau d'Automates Stochastiques

AS	*Autonomous System*, domaine
BGP	*Border Gateway Protocol*
CRM	*Congestion Resolution Mecanism*
CSMA/CA	*Carrier Sense Multiple Access with Collision Avoidance*
FDL	*Fiber Delay Line*, ligne optique à retard
FDMA	*Frequency Division Multiple Access*
GMPLS	*Generalized Multi-Protocol Label Switching*
GSP	*Grid Service Provider*
IP	*Internet Protocol*
LTE	*Long Term Evolution*
MTU	*Maximum Transmission Unit*
MAC	*Medium Access Control*
MAN	*Metropolitan Area Network*
MPLS	*Multiprotocol Label Switching*
OADM	*Optical Add Drop Multiplexer*
OFDMA	*Orthogonal Frequency-Division Multiple Access*
PLR	*Packet Loss Ratio*
PMR	*Private Mobile Radio*
QoS, QdS	*Quality of Service*, Qualité de Service
SC-FDMA	*Single-Carrier FDMA*
SDH	*Synchronous Digital Hierarchy*
SRV	*Scheduling, Reconfiguration and Virtualization*
TDMA	*Time Division Multiple Access*
UMTS	*Universal Mobile Telecommunications System*
WAN	*Wide Area Network*
WiMAX	*Worldwide Interoperability for Microwave Access*
WLAN	*Wireless LAN*

Chapitre I

Introduction

Dans cette partie je retrace mon parcours et présente les thématiques abordées depuis la soutenance de la thèse de doctorat.

Durant ma maîtrise en informatique au sein de l'Ecole Polytechnique Silésienne à Gliwice en Pologne, j'ai assisté aux cours sur l'évaluation de performance du professeur Tadeusz Czachórski. Ses qualités pédagogiques ont rendu ce sujet passionnant à mes yeux. C'est ainsi que l'évaluation de performance est devenue une problématique de recherche sur laquelle je travaille depuis la fin de mes études. A cette époque, la fin des années quatre-vingt, l'analyse de performances portaient principalement sur les architectures d'ordinateurs. Les réseaux locaux étaient déjà déployés mais il manquait toujours l'infrastructure les connectant en réseau global. Le développement rapide et massif des interconnexions entre les réseaux locaux, la construction de réseaux de cœur et métropolitains et, enfin, l'expansion du réseau global Internet couvrant quasiment toute la planète a fait de la problématique de l'efficacité de réseaux un sujet dominant dans le domaine de l'évaluation de performance.

Après la fin de mes études j'ai travaillé à l'Institut d'Informatique Théorique et Appliquée de l'Académie se Sciences de Pologne dans l'équipe dirigée par Tadeusz Czachórski. Les relations bilatérales entre le CNRS et l'Académie se Sciences de Pologne étaient déjà très fortes. Grâce à elles j'ai pu collaborer avec les chercheurs français qui s'intéressaient à la modélisation de réseaux en utilisant une approche analytique. Je parle ici de l'équipe d'évaluation de performance du laboratoire PRiSM de l'Université Saint Quentin en Yvelines à Versailles dirigée par Jean-Michel Fourneau et de l'équipe d'évaluation de performance de l'Institut National des Télécommunications à Evry dirigée à l'époque par Gérard Hebuterne. Les contacts avec l'équipe versaillaise m'ont amenée à m'intéresser aux méthodes compositionnelles de la construction des chaînes de Markov multidimensionnelles telle la méthode des Réseaux d'Automates Stochastiques qui est devenue, plus précisément ses transitions fonctionnelles, le sujet de ma thèse de doctorat.

Ce document résume une partie des thèmes de recherche que j'ai abordés depuis ma soutenance de thèse de doctorat en juillet 1998. Il comporte une étude sur la caractérisation de trafic avec des méthodes markoviennes (1999) réalisée au sein de l'Institut d'Informatique Théorique et Appliquée. Il contient également des travaux effectués à l'INT lors de mes deux séjours de trois mois en 1998/1990 et de six mois en 2002. Ils ont été faits dans le cadre des projets RNRT ROM et ROMEO auxquels participait notamment Alcatel. Il s'agissait de la modélisation d'un commutateur multiservice (en collaboration avec Halima Elbiaze et Tülin Atmaca), d'un commutateur tout optique d'un réseau maillé (avec Ivan Kotuliak et Tülin Atmaca) et du mécanisme d'accès au réseau optique en anneau composé de deux fibres (avec Daniel Popa).

L'intérêt que j'avais développé pour les méthodes compositionnelles de la construction des chaînes de Markov m'a amenée à passer une année (1999/2000) à l'Université d'Edinburgh en Ecosse dans

l'équipe de Fondements d'Informatique. J'y travaillais sur certains aspects d'agrégation de composants de modèles exprimés en termes de l'Algèbre de Processus de l'Evaluation de Performance (PEPA, Performance Evaluation Process Algebra) avec Jane Hillston, qui est inventrice de ce formalisme, dans le cadre du projet scientifique britannique nommé COMPA.

En 2002, pour des raisons personnelles, j'avais quitté la Pologne pour la France où j'ai été aussitôt recrutée comme enseignant-chercheur au sein de l'Ecole Supérieure d'Electricité, Supélec, à Gif sur Yvette dans le département informatique. L'évaluation de performance ne faisait pas partie des axes prioritaires de recherches menés au sein du département. Néanmoins, j'ai eu poursuivre une activité de recherche dans mon domaine de prédilection car les chefs de l'équipe (d'abord Olivier Friedel et, ensuite, Yolaine Bourda) encourageaient mes efforts.

Les contacts avec des industriels ont apporté des financements notamment des bourses de thèses. Elles portent sur les aspects de qualité de service dans des réseaux inter-domaine (Marc-Antoine Weisser) et sur l'interface entre des applications distribuées et le réseau optique reconfigurable à très haut débit (Vincent Reinhard). Ce dernier sujet a été réalisé dans le cadre du deuxième sous-projet du projet CARRIOCAS (Calcul Réparti sur Réseau Internet Optique à Capacité Surmultipliée) faisant partie du Pôle de Compétitive de l'Ile-de-France SYSTEM@TIC. Un autre projet de SYSTEM@TIC, RAF (Réseaux Ad hoc à Forte efficacité), fait conjointement avec IEF ParisSud (Stéphane Pomportes), permet de proposer des modèles de mobilité pour un réseau ad hoc pour des secouristes intervenant sur un lieu d'accident et de travailler sur des algorithmes d'allocation de slots dans ces réseaux avec TDMA.

Ces travaux entamés à Supélec m'ont permis d'intégrer des groupes de travail comme celui consacré aux études des propriétés du protocole de routage des réseaux inter-domaines BGP. Grâce aux contacts avec Alcatel-Lucent France j'ai pu m'investir également dans la recherche sur un nouveau mécanisme de routage dans des réseaux d'accès au réseau radio dont l'architecture est entièrement maillée. Enfin, la participation de Supélec dans la structure scientifique Digiteo regroupant des unités de recherche localisées au sud de Paris, sur le Plateau du Moulon et ses alentours, m'a ouvert la porte au groupe de travail sur les réseaux radio et au groupe qui s'intéresse à l'optimisation au sens large du terme (Optimeo). Digiteo organise annuellement un appel à projets scientifiques et grâce au bon classement du projet DOROthéE (DimensiOnnement des Réseaux métropolitains tout-Optiques à faible consommation Energétique), déposé avec PRiSM lors du concours 2009 concernant le dimensionnement d'un réseau tout-optique, une thèse sur ce sujet est en cours (David Poulain). Je suis donc confiante que la recherche sur l'évaluation de performance de réseaux a de très bonnes perspectives dans les années qui viennent.

Ce mémoire est organisé en six chapitres (y compris celui-ci). Le deuxième chapitre présente les architectures de réseau qui ont été étudies ainsi que les éléments de réseaux qui ont été le sujet de l'analyse. Ce chapitre présente également les propositions faites pour la solution du contrôle d'accès dans des réseaux optiques maillés. Dans le chapitre III nous décrivons des propositions faites afin d'introduire la QdS dans des réseaux inter-domaines, la problématique de transmissions *multicast* (routage et dimensionnement), la réduction du coût d'infrastructure d'un anneau tout-optique et les algorithmes de résolution de conflits d'interférences et d'allocation de ressources dans des réseaux ad hoc ayant une forte demande de QdS. Dans le chapitre IV nous verrons la panoplie des méthodes formelles utilisées pour la modélisation par les chaînes de Markov ainsi que la description de notre générateur de topologies aléatoires dédié à des réseaux inter-domaines avec hiérarchie d'AS et la description d'un modèle de mobilité compositionnel visant modéliser des déplacements des équipes de secouristes sur le lieu de leur intervention. Le chapitre V traite des études de cas de modélisation markovienne des éléments de réseaux. Le dernier chapitre présente les perspectives nos travaux.

Chapitre II

Modélisation d'éléments du réseau

Les réseaux de télécommunications ont ou auront rapidement à faire face à l'augmentation exponentielle du trafic due, d'une part, à l'augmentation continue du nombre d'utilisateurs d'Internet et d'autre part à l'arrivée de services variés à forte demande en qualité ces services (transmissions vidéo, *Voice over IP*, applications distribuées, etc.). Ces services nécessitent la garantie d'une qualité de service (QdS) qui est définie formellement dans [IT] comme l'effet du fonctionnement d'un réseau déterminant le niveau de satisfaction de son utilisateur. Habituellement, la QdS est comprise comme la performance du réseau. Les mécanismes introduits afin de garantir la QdS contribuent à l'amélioration de la performance d'un réseau et peuvent être implémentés à différents niveaux. Par exemple, un mécanisme de gestion de trafic dans un réseau composé de *buffers* dotés de disciplines particulières est un mécanisme de la QdS. Dans le contexte d'un réseau les paramètres de la QdS peuvent notamment être exprimés en termes de débit, délai, fiabilité, séquencement, gigue. Dans le réseau global Internet la QdS de transmissions doit être assurée à tout niveau hiérarchique. Cela signifie qu'elle doit être satisfaite dedans des domaines construisant Internet et pour le réseau inter-domaine.

Dans la suite de ce chapitre nous présentons les éléments de réseau qui ont été modélisés (flots de paquets, routeurs et commutateurs optiques, points d'accès à un réseau optique en anneau). Après ce panel nous décrirons des réseaux qui ont été des sujets de recherche dans leur intégralité : le réseau global inter-domaine et des réseaux optiques maillés d'opérateurs capables de fournir de services à des applications distribuées. La description de ce deuxième type de réseaux est complétée par nos propositions concernant le contrôle d'émission des applications distribuées.

1 Flux de paquets

Les flux de données passant par le réseau d'un opérateur sont très irréguliers et ils manifestent souvent la propriété d'autosimilarité qui signifie que la vitesse de diminution du coefficient d'autocorrélation de ce trafic est faible. Les flux de données passant entre des domaines sont composés d'un grand nombre de flux individuels agrégés. Cette agrégation peut provoquer la perte de la propriété d'autosimilarité du trafic sur les liens inter-domaines. De plus, les mécanismes de gestion d'un réseau comme protocoles génèrent eux-mêmes un certain volume de trafic de signalisation et de contrôle. Ce trafic doit être absolument prioritaire, sa quantité doit être réduite le plus possible et sa diffusion doit être toujours assurée.

L'étude de l'évaluation de performance se déroule dons sur des plusieurs niveaux et traite des aspects différents de la fonctionnalité des réseaux.

Comme cela a été indiqué ci-dessus, dans des réseaux dont la capacité d'acheminement de trafic

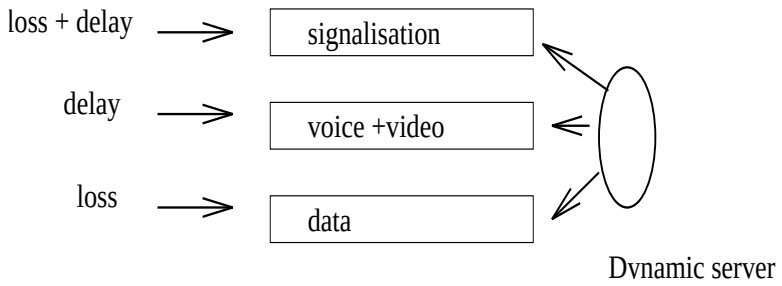

FIG. II.1: *Architecture du commutateur multiservice dans un LAN*

est moyenne, des flots de paquets sont autosimilaires. Cette propriété rend difficile la modélisation d'un tel trafic par le biais des modèles markoviennes. Les sources de trafic poissoniennes génèrent le trafic dont le coefficient d'autocorrelation converge rapidement vers zéro suivant une fonction exponentielle dont l'exposant est négatif. Le coefficient d'autocorrelation d'un trafic autosimilaire tend vers zéro plus lentement en montrant la présence de la mémoire d'une distribution aléatoire le décrivant. Une possibilité de modéliser dans le contexte d'un modèle markovien les sources de trafic dont la vitesse de convergence du coefficient d'autocorrelation à zéro est ralentie par rapport au trafic poissonien est l'utilisation de MMPP (*Markov Modulated Poisson Process*) dont le coefficient d'autocorrelation est représenté par une combinaison linéaire des fonctions exponentielles avec des exposants négatifs. Ce type de sources de trafic a été utilisé pour générer des trafics très variables, notamment dans [RB96]. Nous avons fait une étude sur ce type de sources [Tom99] et nous l'avons largement utilisé en en construisant des modèles markoviens [TEA00, TKA03, TK09]. Les modèles construits avec la méthode des automates stochastiques seront discutés dans les paragraphes 1, 2 et 3.1 du chapitre V.

2 Commutateurs et routeurs

Un réseau de télécommunication assure le service de connectivité par envoi de paquets entre ses éléments intermédiaires comme des commutateurs et des routeurs. Ces éléments effectuent normalement une analyse d'entêtes de paquets arrivants et les renvoient en aval vers un autre élément intermédiaire qui se trouve sur la route vers la destination de paquet. Le rôle joué par un élément assurant le transit peut être décliné en fonction des classes de service auxquelles appartiennent les paquets arrivants. Cette fonctionnalité est donc utilisée pour introduire ces différents traitements de paquets dépendant de la qualité de service demandée par chaque classe et elle constitue l'élément essentiel de la gestion de la qualité de service.

La nécessité de la diversification des traitements de paquets apparaît dans des réseaux conventionnels et optiques à des niveaux différents de la hiérarchie du réseau global (LAN, MAN, WAN). La proposition d'un mécanisme de diversification de traitement de paquets dans un commutateur de LAN d'entreprise privée dont l'architecture est schématisée sur la Figure II.1 par une file prioritaire ainsi que son modèle markovien seront présentés dans le paragraphe 1 du chapitre V.

Les avancées de la technologie optique créent un enjeu important des nouvelles architectures de commutateurs. Dans les architectures "anciennes" l'analyse de l'entête d'un paquet optique se fait après sa conversion en forme électronique. Après la récupération de l'entête un paquet est de nouveau transformé en forme optique. Les nouvelles architectures n'introduisent pas le délai dû par ces conversions car tout le traitement de paquet est effectué sur la couche optique [GR03] mais elles

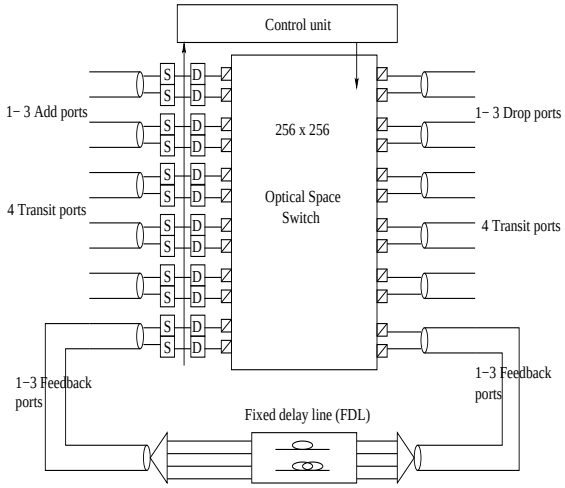

Fig. II.2: *Architecture du commutateur optique étudié*

doivent tenir compte de nombreuses contraintes technologiques, notamment l'absence de tampons optiques.

Selon [C+06] il y a deux approches opposées envisageables dans le déploiement de WAN tout optiques : soit les paquets optiques sont de taille fixe et leur traitement est synchronisé [G+01], soit les paquets sont de taille variable et la commutation est basée sur le principe de *burst switching* [Qia00]. Les études que nous avons effectuées [TKA03, TK09] concernent le premier de ces deux paradigmes car il permet de garder le même format de paquet [HA00]. Les résultats de nos travaux ont été confrontés à ceux du projet ROM [G+01]. La nature synchronisée du réseau étudié détermine l'espace du temps discret de la chaîne de Markov que nous utilisons pour modéliser un commutateur.

L'architecture du commutateur optique que nous avons étudié est représenté sur Figure II.2. Les paquets entrant dans le commutateur sur plusieurs longueurs d'onde par fibre sont synchronisés (S). Ensuite les entêtes sont enlevées afin d'effectuer leur traitement électronique. Les données (*payload*), toujours sous la forme optique, transitent par les lignes à retard (D) (*Fiber Delay Line*, FDL) [TTC99] et sont ensuite dirigées vers la matrice du commutateur (*non-blocking switching matrix*). Un délai introduit par FDLs est nécessaire pour que l'unité de contrôle (CU) ait le temps pour analyser des entêtes et configurer la matrice de commutateur pour qu'un paquet puisse la traverser et en sortir sur un port et une longueur d'onde choisis. Avant de quitter le commutateur, un paquet reçoit une nouvelle entête. Le traitement effectué dure typiquement quelques centaines de nanosecondes.

Un problème majeur à résoudre dans le cas d'un commutateur est la réduction de congestion qui représente une cause de pertes de paquets. Nous avons étudié trois architectures de commutateur différentes par rapport au mécanisme de résolution de la congestion (CRM, *Congestion Resolution Mechanism*) utilisé.

1. Aucun mécanisme de résolution : l'architecture du commutateur est identique à celle décrite ci-dessous et les pertes se manifestent dans des cas où au moins deux paquets sont dirigés vers le même port de sortie. Nous étudions cette architecture car elle est la moins coûteuse bien qu'elle ne prévienne pas les pertes.
2. Mémoire : dans la technologie optique la mémoire est émulée par les lignes à retard. Des données entrant dans le commutateur sont dirigées vers des FDLs introduisant des délais différents. L'unité centrale peut donc retarder les paquets T, $T + k$, $T + 2k$, ..., $T + nk$ où

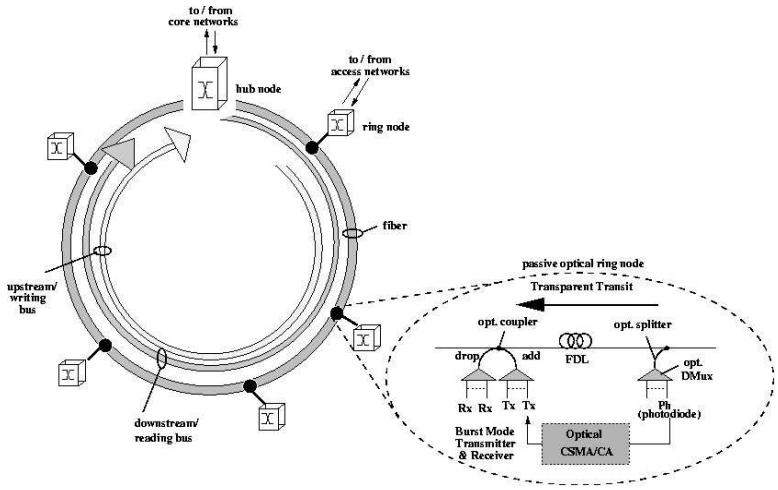

Fig. II.3: *Architecture du réseau optique DBORN*

T indique le temps nécessaire pour le traitement d'un paquet et k est une unité de retard habituellement égale à la longueur du paquet. Si un paquet ne trouve pas de port de sortie au bout de n retards, il est détruit.

3. *Feed-back* : des paquets qui ne trouvent pas un port libre vers la destination sont redirigés vers un port d'entrée en subissant un délai. Cette architecture est relativement simple à réaliser mais elle peut introduire le déséquencement de paquets. Malgré cette propriété non souhaitable pour des transmissions de paquets IP, cette solution peut être considérée comme un bon compromis entre le prix d'équipement et sa performance.

L'étude du cas concernant la modélisation par les chaînes de Markov en temps discret d'un commutateur optique est présentée dans le paragraphe 2 du chapitre V.

3 Points d'accès au réseau optique

Un aspect principal de l'architecture de réseau métropolitain optique est un mécanisme d'accès à fibre efficace et flexible [CWK00, O'M01]. Nous avons travaillé sur cet axe en prenant comme architecture de référence DBORN (*Dual-Bus Optical Ring Network*) [BBDP05] élaborée par Alcatel qui est représentée dans la Figure II.3. Nous l'avons étudiée dans deux cas opposés : asynchrone (des paquets de taille variable) et synchrone (des paquets de taille fixe).

L'approche CSMA/CA (*Carrier Sense Multiple Access with Collision Avoidance*) asynchrone a été proposée dans des architectures expérimentales comme DBORN mentionnée ci-dessus et HORNET (*Hybrid Optical Ring NETwork*) [WRSK03].

Il y a plusieurs raisons pour lesquelles cette approche apparaît adaptée comme protocole d'accès dans des réseaux métropolitains. Premièrement, elle permet une solution totalement synchronisée, sans contrôleur synchronisé ni longueurs d'onde dédiées à la synchronisation. Deuxièmement, elle permet de partager efficacement la bande passante entre deux points d'accès à l'échelle du temps d'un paquet. Troisièmement, une configuration optique passive avec CSMA/CA permet de réduire le coût d'équipement [SDD+01]. Les inconvénients de cette approche asynchrone que nous pouvons noter sont la priorité dépendant de la localisation (*positional priority*) [CO93] (des nœuds proches du concentrateur ont l'accès au médium physique plus facile que celles qui en sont éloignées) et la

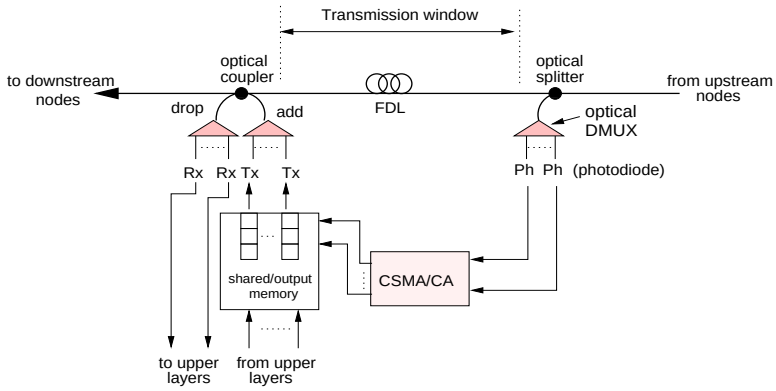

Fig. II.4: *Proposition de l'interface de accès au réseau DBORN asynchrone*

fragmentation de la bande passante (*bandwidth fragmentation*) [PA06] se manifestant par l'apparition sur la fibre de périodes vides mais courtes et, par conséquent, inutilisables.

Les éléments caractéristiques de l'architecture DBORN sont : deux bus optiques et le concentrateur (*hub*), le seul nœud dans lequel la conversion O/E/O peut avoir lieu. Les nœuds du réseau sont équipés de multiplexeurs transparents et passifs dits *Optical Add Drop Multiplexers* (OADMs) composés de *transmitters* (Tx) réalisant la fonctionnalité ADD et de *receivers* (Rx) réalisant la fonctionnalité DROP. Deux bus sont réalisés par la séparation spectrale de longueurs d'onde de la fibre dans le but de repartir des actions de lecture et l'écriture. Le bus descendant connectant le concentrateur aux nœuds (*downstream bus*) effectue les transmissions point-multipoint pour la lecture des données par les nœuds et peut être réalisé avec la technologie synchronique comme SDH/SONET. Le bus montant passant entre les nœuds et le concentrateur (*upstream bus*) assure les transmissions multipoint-point pour l'insertion de paquets dont la technologie est basée sur le mode asynchrone (*Burst Mode Transmitters -BMTx-* dans les nœuds et *Burst Mode Receivers -BMRx-* dans le concentrateur) et sur la structuration de la deuxième couche selon Ethernet.

Les multiplexeurs OADM présents dans les nœuds insèrent/extraient (ADD/DROP) des paquets dans/des longueurs d'onde communes. L'écriture des paquets destinés au concentrateur se fait sur le *upstream bus* et la lecture des paquets transmis par le concentrateur se fait sur le *downstream bus*. Par conséquent, les nœuds ne peuvent pas communiquer directement mais uniquement par intermédiaire du concentrateur et la construction d'OADMs est simplifiée. Ces multiplexeurs ne contiennent pas de tampons optiques et pour cette raison les paquets injectés dans la fibre ne subissent pas de pertes. Par contre, le contrôle d'accès dans un nœud devient indispensable. Une réalisation possible du protocole d'accès basé sur CSMA/CA avec une photodiode introduisant une fenêtre glissante (*sliding window*) et FDLs est illustrée par la Figure II.4. Cette implémentation permet à un nœud de détecter de créneaux disponibles. Si un paquet électronique est plus long qu'un créneau disponible, il reste stocké dans le nœud. Le retard induit par les FDLs est légèrement plus grand que la plus grande longueur possible d'un paquet dans Ethernet (*Maximum Transmission Unit, MTU*) afin d'éviter des collisions.

Les réseaux optiques asynchrones ont été peu étudiés dans le contexte de leur performance par le biais de modèles analytiques. Nous pouvons mentionner ici [BBDP05] et [CH04]. La difficulté principale de modélisation analytique d'un réseau à paquets de taille variable provient des phénomènes de fragmentation de la bande passante et du blocage HoL (*Head of Line*) [PA06]. Le blocage HoL se produit dans une file d'attente de discipline FIFO quand des paquets longs se trouvant en tête ne peuvent pas être insérés dans la fibre car les créneaux disponibles sont trop courts pour le

faire. Dans cette situation des paquets plus courts qui auraient pu trouver leurs créneaux sur la fibre restent bloqués derrière des paquets longs. A cause du phénomène HoL, des modèles markoviens avec la discipline FIFO probabiliste produisent des résultats trop optimistes. Le modèle markovien représenté dans le formalisme de Réseaux d'Automates Stochastiques du réseau de l'architecture DBORN que nous avons proposé sera décrit dans le paragraphe 3.1 du chapitre V.

Les inconvénients de l'approche synchrone [YDM00] sont provoqués par le fait que le taux de remplissage d'un paquet de taille fixe peut être faible (cet espace libre est inutilisable) [Pa03] et — dualement — l'attente trop longue par des paquets de la formation d'un paquet optique complet les retarde d'une manière inacceptable d'un point de vue de la qualité de service demandée. Ce deuxième problème peut être résolu par l'introduction d'un mécanisme temporisateur *Time-Out* [KAP03]. La construction d'une chaîne de Markov en temps discret pour modéliser un point d'accès à un réseau DBORN sera présentée dans le paragraphe 3.2 du chapitre V.

4 Réseaux complets

L'analyse détaillée des éléments d'un réseau dans le contexte de leur performance n'est pas suffisante pour répondre aux exigences relatives à la qualité de service requise par les utilisateurs de réseaux et leurs opérateurs. Un utilisateur d'un réseau est intéressé de savoir si ses communications se passeront avec la qualité de service définie par le contrat pour lequel il paie. D'un autre côté, un opérateur veut savoir comment maximiser son gain financier en vendant des contrats et en utilisant le mieux possible l'infrastructure existante de son réseau. Un fournisseur de services de réseau peut aussi vouloir connaître l'ampleur du développement de son réseau quand il souhaite élargir son offre commerciale à ses clients. Il devient donc indispensable de mener l'analyse d'un réseau complet pour obtenir des informations concernant la qualité de service pour des transmissions en bout en bout. Nous avons travaillé sur ce thème dans le contexte du réseau inter-domaine Internet et d'un réseau optique commercial à très grand débit offrant des services pour des applications grilles.

4.1 Réseau inter-domaine Internet

Le réseau Internet, vu au niveau global, est un réseau composé de réseaux indépendants appartenant à différents opérateurs. Ces réseaux, nommés domaines, utilisent tous un protocole de routage BGP (*Border Gateway Protocol*) [RLH06] pour pouvoir établir des routes. Ce protocole fonctionne selon le principe du meilleur effort (*Best Effort*). Les domaines ne disposent donc d'aucune information sur la qualité de service des routes connues. En contrepartie, les opérateurs de domaines gardent la confidentialité de la configuration de leurs infrastructures et leurs performances. La possibilité de maîtriser la qualité de service dans le réseau inter-domaine exige de faire un compromis entre la préservation des informations confidentielles et une diffusion des informations permettant d'établir des routes satisfaisant certains critères de qualité de service.

Après avoir envisagé trois options pour introduire l'aspect de la qualité de service dans le réseaux inter-domaines (remplacer BGP, améliorer BGP, fonctionner en parallèle de BGP), nous avons opté pour cette dernière approche qui se superpose à BGP. Elle a été ensuite abordée selon deux perspectives différentes : établir de routes avec les garanties de qualité strictes (approche connectée, paradigme IntServ [BCS09]) et fluidifier le passage de trafic avec exigences de la qualité de service sans les garantir au sens propre du terme (approche déconnectée, paradigme DiffServ [BBC+98]).

La solution algorithmique proposée pour cette deuxième approche se base sur l'observation qu'il existe une hiérarchie dans le réseau Internet introduite par les liens établis entre des domaines. Ces liens reposent sur des contrats. D'un point de vue économique, les liens peuvent être regroupés en trois types [Gao01, SF04] :

- les liens de client à fournisseur (*customer-to-provider, C2P*) qui sont établis entre un domaine client qui achète de la capacité de transit vers une partie du réseau à un autre domaine, le fournisseur ;
- les liens de pair à pair (*peer-to-peer, P2P*) qui sont établis entre des domaines de taille comparable qui échangent gratuitement de leur capacité de transit ;
- enfin d'autres liens existants comme les liens *S2S*, de *sibling* à *sibling*, qui relient deux domaines appartenant au même opérateur.

L'existence de ces relations économiques a une forte influence sur les décisions prises arbitrairement par des domaines concernant les informations diffusées par BGP sur les routes annoncées et sur les destinataires de ces messages. Aucun domaine n'a intérêt à informer les autres des destinations vers lesquelles ils peuvent faire transiter le trafic quand ils doivent payer pour ses transmissions. En général, les pairs s'informent mutuellement des routes vers leurs clients. Les fournisseurs annoncent toutes les routes à leurs clients et les clients envoient des annonces à fournisseurs et pairs mais aussi à leurs propres clients. La liberté des domaines de choisir une route à annoncer et les relations commerciales entre des domaines font que les routes du réseau inter-domaines ne sont nécessairement pas des plus courts chemins [GW02]. Les routes ont une forme particulière appelée par Gao [Gao01] *sans-vallée* (Définition II.1).

Définition II.1 *Une route* sans-vallée *est constituée de trois parties successives :*
- *une suite (éventuellement vide) composée exclusivement de liens de* C2P*,*
- *zéro ou un lien* P2P*,*
- *une suite (éventuellement vide) exclusivement composée de liens* C2P *inversés.*

Cette hiérarchie, bien qu'elle soit connue des scientifiques et ingénieurs, n'a jamais été utilisé pour l'envoi de messages protocolaires. Nous avons donc été confrontés au problème du manque des modèles de réseaux inter-domaines dotés de cette hiérarchie. Pour y remédier, nous avons conçu un algorithme permettant d'induire la hiérarchie inter-domaine dans un modèle de topologie inter-domaine plate et l'avons implémenté.

L'introduction de la hiérarchie dans un topologie aléatoire plate est présentée dans [WT06b, WT07a, TW10a] (le paragraphe 2 du chapitre IV de ce mémoire). L'approche déconnectée [WT07b, WTB08, WTB10] est décrite dans le paragraphe 1.1 du chapitre III et elle est suivie par l'approchée connectée [WT05, WT06a]. La problématique dans son intégralité de la satisfaction de la qualité de service dans le réseau inter-domaine est discuté dans [Wei07].

4.2 Contrôle d'émission dans un réseau offrant des services pour des grilles

Le projet CARRIOCAS (Calcul Réparti sur Réseau Internet Optique à Capacité Surmultipliée) [Ver07, CAR, VAB+08] du Pôle de Compétitivité SYSTEM@TIC auquel nous avons participé concernait la conception architecturale d'un réseau commercial qui est adapté à fournir à ses clients non seulement un service de connectivité mais aussi des services avancés comme des capacités de calcul et de stockage. Un tel réseau peut donc servir de support à différentes applications distribuées de type grille (*grid*) [FK01, FK02] qui jusqu'à présent utilisent leurs propres infrastructures de réseau dans lesquelles aucun autre trafic ne circule. Les programmes *middle-ware* [FK97, Fos05] implantent un *work-flow* d'une application sur un réseau sans tenir compte de la disponibilité de la bande passante.

La Figure II.5 schématise les éléments principaux de l'architecture CARRIOCAS [Aud07, CAR]. L'allocation des ressources de réseau est effectuée par l'unité centrale SRV (*Scheduling, Reconfiguration and Virtualisation*) gérée par l'opérateur du réseau. Les unités appelées *Data Centers* fournissent des services particuliers pour les grilles. Le SRV se charge d'établir des connections entre

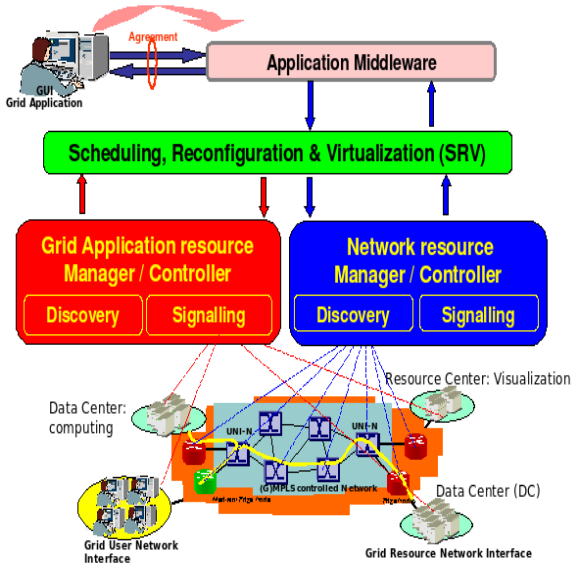

Fig. II.5: *Proposition de l'architecture CARRIOCAS d'un réseau optique*

eux selon les besoins d'application exprimés par un utilisateur. Un utilisateur souhaitant lancer une application distribuée s'adresse aux unités GSP (*Grid Service Provider*) qui formule d'une demande de réservation de services de connectivité (une demande faite au réseau), de stockage et de calcul (auprès de *Data Centers*) en respectant ses besoins décrits sous la forme d'un *work-flow*, et lui proposent un prix. Les unités GSPs ne sont pas nécessairement gérées par l'opérateur du réseau mais ils peuvent appartenir à d'autres organismes commerciaux. Cette séparation de services de connectivité et de services distribués propres aux grilles encourage la concurrence et est recommandée par les institutions européennes régulatrices du marché de télécommunications.

L'aspect novateur de l'architecture CARRIOCAS consiste à pouvoir gérer l'accès des applications distribuées aux ressources de réseaux. Il impose la proposition de mécanismes pouvant réaliser ce contrôle et d'algorithmes permettant d'établir des réservations pour les grilles respectant à la fois les demandes des utilisateurs et optimisant la quantité de ressources de réseau qui y sont engagées.

Dans un premier temps nous nous sommes intéressés à la réactivité du réseau vis-à-vis les applications, sujet qui est discuté dans le paragraphe suivant, et nous avons proposé un mécanisme d'allocation de la bande passante avec contrôle d'émission de grilles [RT08b, RT08a]. Ensuite nous nous sommes tournés vers la problématique du déploiement de l'infrastructure du réseau, plus précisément, du déploiement de l'infrastructure de routeurs optiques permettant la réplication de transmissions pour réaliser des connections *multicast* [RTBW09, RCT$^+$11b, RCT$^+$11a] (cf. le paragraphe 2 du chapitre III).

Dans [RT08b, RT08a] nos avons proposé un mécanisme du contrôle d'émission centralisé respectant les contraintes architecturales exprimées par la suite d'interactions suivante :

- Le SRV récupère des demandes envoyées par le GSP et reçoit des informations concernant la bande passante disponible et la latence estimée. Il choisit des ressources de réseau correspondantes à chaque demande et calcule leur prix. Ensuite il communique le prix au GSP et réserve les ressources dans le cas où le GSP accepte cette offre.
- Le GSP permet à un utilisateur de fournir la description de ses applications et leurs besoins. Cet élément est également responsable de la sélection des ressources autres que les ressources

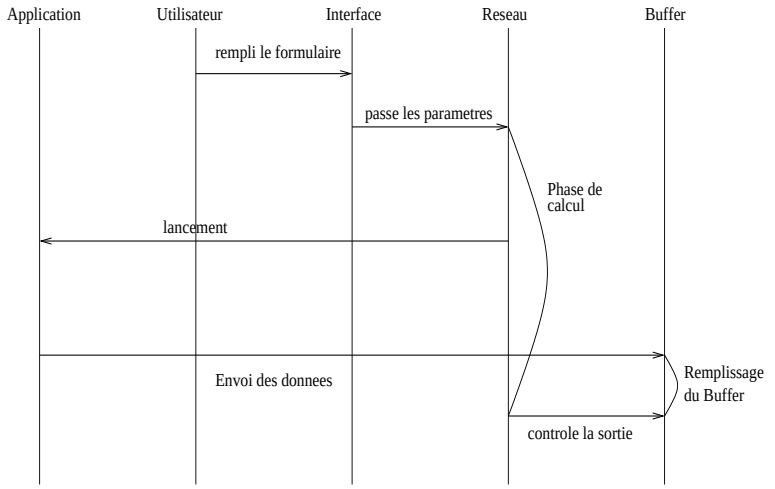

Fig. II.6: *Scénario décrivant une session de lancement d'une application distribuée sous la condition que les ressources demandées soient disponibles*

de réseau et de leur réservation. Il est chargé des négociations concernant les ressources de réseau avec le SRV.

- Selon l'architecture de CARRIOCAS, le SRV obtient des informations concernant les ressources de réseau du gestionnaire de ressources de réseau qui observe de l'état du réseau.

Dans cette optique vous avons proposé un mécanisme de contrôle basé sur des tampons attribuant un tampon pour toute application dont le fonctionnement est décrit comme suit :

- Un GSP crée un tampon et calcule ses paramètres en fonction des besoins de l'application cible.
- L'application commence à envoyer des données dans son tampon après avoir reçu le signal de confirmation émis par le GSP.
- Le tampon reçoit des données de l'application et les retransmet dans le réseau.
- Le taux d'envoi de données depuis le tampon est contrôlé par le SRV allouant de la bande passante disponible pour des applications.

Le scénario que nous avons proposé (Figure II.6) présente des interactions entre un utilisateur, son interface, le tampon de son application, un GSP et le réseau lors du lancement d'une application distribuée. Avant d'exécuter l'application, l'utilisateur décrit au GSP ses besoins en termes de *work-flow*, de latence maximale et de bande passante maximale. Le GSP s'occupe lui-même de la réservation des ressources de calcul et de stockage. Pour réaliser la réservation de ressources du réseau, le GSP s'adresse au SRV qui vérifie la disponibilité des ressources demandées et propose un prix. Si le prix est accepté par le GSP, le SRV alloue les ressources et notifie le GSP. Par la suite le GSP envoie le signal de confirmation (ACK) à l'application qui commence aussitôt à transmettre ses données dans le tampon. En même temps le réseau est reconfiguré pour satisfaire la demande de l'utilisateur. Après la reconfiguration, le SRV active la sortie du tampon de l'application.

Les simulations ont été effectuées pour les applications distribuées massives de visualisation à distance dont la spécification avait été fournie par EDF [EDF]. Nous avons notamment étudié l'impact sur le temps d'attente dans des tampons et sur les pertes du lancement de plusieurs applications demandant une QdS. Nous avons ainsi mesuré l'importance du délai de changement d'état et du délai de notification de changement d'état. Par état du réseau nous entendons ici la bande passante disponible. Le délai de changement d'état est alors une période entre deux changements

de volumes de bande passante disponible. Le deuxième délai est une période entre un changement d'état du réseau et l'ouverture d'un tampon. Les paramètres de performance considérés sont le temps moyen d'attente pour émettre et le taux de pertes de données. Les résultats commentés en détail dans [RT08b, RT08a] montrent l'influence très forte de ces deux délais sur la performance du réseau. Le réseau devient difficile à contrôler dans le cas où le délai de notification est plus long que le délai de changement d'état. Ces observations permettent aux concepteurs du réseau CARRIOCAS de tenir compte de l'importance du *monitoring* du réseau.

Chapitre III

Amélioration des performances de réseaux

Les services fournis par un réseau et leurs propriétés font l'objet d'un contrat. Ce contrat précise la qualité de service demandée par un client qu'un opérateur de réseau s'engage à respecter. L'opérateur de réseau voulant fournir des services de la qualité demandée par ses clients souhaite cependant rentabiliser le plus possible l'infrastructure existante de son réseau ou bien faire un projet du déploiement d'une nouvelle infrastructure dont le coût de construction est le moins élevé possible. Le problème de la garantie de la qualité de service peut se heurter à un manque d'information nécessaire pour estimer la qualité de routes ou de mécanisme pour annoncer la classe de service des paquets. Nous avons traité la problématique de la garantie de la qualité de service dans de différents réseaux dont les contraintes technologiques déterminent l'approche algorithmique utilisée.

Le paragraphe 1 est consacré à la présentation des mécanismes proposés afin d'introduire des garanties de qualité de service dans le réseau inter-domaine Internet. Le paragraphe suivant traite l'utilisation optimale de ressources dans un réseau optique maillé dont l'architecture correspond à celle du réseau CARRIOCAS (cf. le paragraphe 4.2 du chapitre II). Le paragraphe 3 continue la présentation de nos travaux concernant des réseaux optiques. Ils sont consacrés au dimensionnement d'un réseau métropolitain dans les nœuds duquel les transmetteurs (Tx) et récepteurs (Rx) sont séparés. Le paragraphe 4 de ce chapitre aborde la problématique de la garantie de la QdS dans des réseaux ad hoc utilisés par des secouristes dans le cas de leur intervention sur un terrain d'accident. La nature d'intervention sur le lieu d'une catastrophe impose le traitement particulier de ce problème.

1 Qualité de service dans le réseau inter-domaine

La spécificité du réseau inter-domaine Internet est décrite dans le paragraphe 1 du chapitre II. Le protocole BGP assurant le routage dans ce réseau ne fourni de moyens ni pour découvrir la qualité de routes ni pour allouer des ressources. Dans ce paragraphe nous présentons les idées concernant la gestion de la qualité de service en nous servant de nouveaux protocoles collaborant avec BGP. Nous analysons d'abord l'approche sans établissement de connexion (l'approche DiffServ) et, ensuite, l'approche avec établissement de connexion (l'approche IntServ) entre la source et la destination de transmissions.

1.1 Détection de routes congestionnées

Vue la nature hétérogène du réseau global de domaines, l'établissement de routes satisfaisant des exigences au niveau de la qualité de service se heurte à la confidentialité de la configuration et de la performance des domaines. Nous avons remarqué que, malgré la préservation de la confidentialité, des domaines fournisseurs sont obligés par des contrats commerciaux de fournir la qualité de service pour laquelle paient leur clients. Autrement dit, des fournisseurs de connectivité ne sont pas intéressés par "tricher" vis-à-vis à leurs clients pour ne pas les perdre au profit de domaines concurants.

Cette observation nous a apporté l'idée que des clients peuvent avoir la connaissance d'atteignabilité de leurs fournisseurs car les fournisseurs sont intéressés par informer leurs clients de leur capacité d'acheminer le trafic. Cette idée nous est apparue d'autant plus intéressante que la diffusion de messages contenant l'information sur la capacité de passer le trafic pourrait être restreinte à une partie du réseau global grâce à la structure hiérarchique de ce réseau. Les messages alertant de problèmes avec l'acheminement de trafic sont envoyés indépendamment du protocole BGP. Notre algorithme utilise les routes récupérées par BGP en choisissant celles qui ne sont pas congestionnées. L'utilisation d'un mécanisme très "léger" et sa possibilité d'agrégation de flux de trafic positionne notre proposition dans le cadre du paradigme DiffServ [BBC+98].

Pour formaliser le problème de la diffusion de messages d'alerte nous représentons le réseau comme un graphe dont les nœuds (l'ensemble V) sont des domaines, les relations *C2P* sont des arcs (l'ensemble A) et les relations *P2P* sont des arêtes (l'ensemble E). Le graphe $G = (V, A, E)$ construit de cette manière est donc un graphe mixte au sens de la définition donnée dans [RV99]. L'existence de relations commerciales implique que ce graphe est sans circuits [GR00, GGR01]. La diffusion d'alertes dépend du routage et du trafic traversant des domaines, alors le graphe G est doté

- d'une fonction de capacité de domaines $c : V \longrightarrow \mathbb{R}_+^*$,
- d'une matrice de trafic $T : V^2 \longrightarrow \mathbb{R}_+$ dont les éléments indiquent le trafic à acheminer le mieux possible et un élément $T_{i,i}$ représente le trafic interne d'un domaine i et
- d'une matrice de routage *sans-vallée* (voir Définition II.1) $R : V^2 \longrightarrow \{1, \ldots, n\}$ dont les valeurs correspondent un *next hop* pour le trafic transitant par le domaine i et dont la destination est le domaine j ($R_{i,i} = i$ pour le trafic interne d'un domaine i).

Nous disons qu'un domaine i est *perturbé* si la quantité totale de trafic transitant par i, émis par i et à destination de i est strictement supérieure à sa capacité. Nous nous focalisons sur la recherche du routage qui minimise le nombre de domaines perturbés car notre algorithme fait envoyer des alertes par des domaines perturbés. Nous espérons qu'en faisant ainsi nous minimisons également la quantité de trafic perturbé (le trafic de i à j est un trafic perturbé si la route allant de i à j contenue dans le routage est perturbée ; une route est perturbée si au moins un des nœuds qui la composent est perturbé).

Le raisonnement détaillé ci-dessus nous mène à poser *le problème du routage sans-vallée* ayant comme but trouver le routage pour lequel au plus K domaines sont perturbés :

Problème III.1 Le problème du routage sans-vallée

Données : G un graphe du réseau inter-domaine, c une fonction de capacité, T une matrice de trafic et K un nombre naturel positif.

Question : Existe-t-il une matrice de routage sans-vallée R tel que le nombre de nœuds perturbés pour la matrice de trafic T est inférieur ou égal à K ?

Nous avons démontré dans [WTB08] que ce problème peut être réduit au problème de Steiner dans les graphes orientés acycliques qui est *NP*-complet [HTWL96]. Nous avons également démontré que la connaissance de solutions pour des instances proches ne permet pas de le ramener à un problème polynomial. Notre rapport [WT07b] contient la preuve que le problème de routage est un

problème inapproximable car il permet de résoudre le problème de couverture d'ensemble, lui même inapproximable [Joh73].

Pour résoudre ce problème nous avons proposé un algorithme heuristique distribué [WTB08, WTB10] fondé sur l'envoi de messages d'alerte transportant les informations concernant la perturbation des domaines. Chaque nœud informe son voisinage (clients, pairs et fournisseurs) lorsqu'il devient perturbé afin que ses voisins puissent adapter leur routage. Chaque nœud informe également ces voisins lorsqu'il retourne dans un état opérationnel. Afin d'éviter de noyer l'intégralité du réseau avec des messages de contrôle, nous utilisons la hiérarchie du réseau inter-domaine pour limiter la portée de la diffusion d'une information.

Notre algorithme se base sur le principe que chaque domaine peut être dans l'un de deux états suivants : vert ou rouge. Un nœud est dans l'état rouge si au moins une des deux conditions suivantes est satisfaite :

- la quantité totale de trafic transitant à travers de ce nœud, à destination de ce nœud ou émis depuis ce nœud est supérieure à sa capacité ;
- au moins un des *next-hops* qui est un fournisseur est dans l'état rouge.

Un nœud qui n'est pas rouge est vert. Un nœud ne devient jamais rouge à cause de l'état de ses pairs ou de ses clients. En conséquence, un nœud appartenant à une couche basse de la hiérarchie et devenant rouge n'aura pas d'impact sur l'ensemble des nœuds.

Pour assurer le fonctionnement de notre algorithme chaque nœud est équipé d'une table qui contient l'état de ses voisins, d'une table contenant pour chaque destination la liste des différents *next-hops* qui peuvent être utilisés pour atteindre chaque destination selon des messages d'annonces de routes BGP et d'une table de routage supplémentaire qui ne contient que les routes qui sont potentiellement non-congestionnées. Pour un nœud donné, si aucun de ses voisins n'est perturbé alors la table de routage prioritaire est la même que la table de routage de BGP.

La performance de l'algorithme proposé a été évaluée en termes du nombre de nœuds perturbés, de la quantité de trafic perturbé et du nombre de messages d'alertes envoyés par simulation avec l'aide du générateur de topologies du réseau inter-domaine avec hiérarchie SHIIP écrit au préalable (voir le paragraphe 2 du chapitre IV). Ces résultats ont été comparés à ceux obtenus avec BGP "pur" fournissant une borne supérieure du nombre d'éléments perturbés et avec BGP doté d'un mécanisme heuristique mais centralisé fonctionnant avec la matrice de trafic fixe que nous avons proposé [WTB08, WTB10] dans le but d'avoir une borne inférieure du nombre d'éléments perturbés. Nous avons observé que notre algorithme améliore significativement l'écoulement du trafic dans le réseau par rapport à BGP "pur" (uniquement dans le cas de la saturation du réseau le nombre de nœuds perturbés s'approche de celui de BGP "pur"). Dans toutes le expériences effectuées notre algorithme convergeait vers la borne minimum. Le nombre de messages de signalisation restait stable durant le temps de simulations. Les résultats obtenus sont très encouragent d'autant plus que cet algorithme se contente d'un nombre d'information de contrôle très réduit.

1.2 Routes multi-critères avec réservation

L'écoulement de trafic prioritaire dont les garanties sur la qualité de service sont strictes ne peut pas être réalisé par l'algorithme détectant les routes potentiellement congestionnées décrit ci-dessus. Le contexte du trafic exigeant la garantie stricte implique la réservation d'une route qui respecte les paramètres de qualité de service demandés comme cela est fait au niveau d'un seul domaine par le protocole RSVP [BZB+97]. L'établissement d'un circuit virtuel inter-domaine réalisé par le protocole MPLS [RVC01] au niveau d'un domaine ou au niveau inter-domaine [FAV08] fait avec les informations diffusées par le protocole BGP, ne tient donc pas compte de paramètres de qualité de service sur les routes annoncées.

La nature des paramètres de la qualité de service comme la bande passante, le délai, le taux de pertes ou la gigue rend difficile le choix de routes car les contraintes additives/multiplicatives doivent être évaluées le long de la route toute entière. Or ces routes traversent des domaines indépendants qui ne révèlent pas les informations précises sur leurs performances. Cela rend la garantie de la qualité de service particulièrement difficile.

Dans un premier temps nous avons formalisé le problème en utilisant des graphes non-orientés (sans hiérarchie). Les garanties sur K paramètres de trafic sur lesquels les domaines donnent les garanties à leurs voisins sont représentés par K fonctions de pondérations attribuées aux arêtes, $w_i : V \longrightarrow \mathbb{R}$, $i \in \{1, 2, \ldots, K\}$. Les paramètres caractéristiques d'acheminement de trafic offerts sont soit additives (délai), multiplicatives (taux de pertes), convexes (bande passante). Une fonction de pondération d'un chemin dans le réseau est donc calculée en tenant compte de ces paramètres. Enfin, nous définissons les prédicats, dont la forme dépend également de la nature des paramètres, indiquant si une contrainte c concernant le paramètre $i, i \in \{1, 2, \ldots, K\}$ est satisfaite pour un chemin p : $S_i : \text{chemin}(G) \times \mathbb{N} \longrightarrow \{\text{vrai}, \text{faux}\}$. Après ces précisions, nous posons formellement le problème à résoudre :

Problème III.2 Le problème de la recherche de chemins multi-contraints

Données : $G = (V, A)$ un graphe non-orienté, w_i un ensemble de fonctions de pondération de paramètres de domaines, W_i un ensemble de fonctions de pondération sur les routes, S_i un ensemble de fonctions de satisfaction de contraintes, $v_o, v_d \in V$ une origine et une destination d'une route et $c_i \in \mathbb{N}$ un ensemble de K contraintes.
Question : Existe-t-il une route p de v_o à v_d pour laquelle $\forall i \in \{1, 2, \ldots, K\}, S_i(W_i(p), c_i) = \text{vrai}$?

Ce problème est *NP*-complet quand le nombre de contraintes additives et multiplicatives est supérieur ou égal à deux [GJ79] (preuve dans [WC96]).

Après avoir déterminé la complexité du problème pour les graphes non-orientés, nous nous sommes posé la question de la complexité de la recherche d'un chemin avec contraintes multiples pour les graphes mixtes avec hiérarchie donc les seules routes permises ont la propriété sans-vallée. Pour ce faire nous commençons par poser le problème suivant :

Problème III.3 Le problème de la recherche de chemin multi-contraint dans les DAG

Données : $G = (V, A)$ un graphe orienté acyclique, $c_1, c_2 : V \longrightarrow \mathbb{R}^+$ deux pondérations c_1 et c_2 sur les éléments de V, o et d deux sommets, K_1 et K_2 deux nombres naturels.
Question : Existe-t-il un chemin p de o à d tel que

$$\begin{cases} \sum_{i \in p} c_1(i) & \leq & K_1 \\ \sum_{i \in p} c_2(i) & \leq & K_2 \end{cases} ?$$

Ce problème est *NP*-complet car nous ont pu démontrer son équivalence avec le problème de partition d'ensemble qui est *NP*-complet [Kar72].

Pour proposer un algorithme heuristique résolvant nos problèmes nous avons étudié les solutions existantes. Celles centralisées comme dans [Jaf84, WC96, GO97, MS97, CN98b] sont inadaptées pour le réseau inter-domaine dans lequel une entité ayant la vision de la totalité du réseau ne peut pas exister. Nous nous sommes donc tournés vers les algorithmes distribués [CN98a] étant conscient que les information sur le réseau diffusées par BGP sont insuffisantes pour les faire fonctionner. La récupération de données supplémentaires ou leur prédiction pour les avoir toutes et suffisamment fréquemment est une tâche très lourde [XLWN02].

Les conclusions auxquelles nous avons abouti nous ont amené à baser notre solution sur une recherche en profondeur intégrant des mécanismes de coupes de branches, à reprendre en partie les idées de [KK01] et à introduire une signalisation consistant à envoyer un message-sonde explorant le réseau afin de trouver une route satisfaisant les contraintes données [WT05, WT06a]. Une sonde est envoyée dans le réseau inter-domaine pour trouver un chemin satisfaisant les contraintes et pour réserver en même temps les ressources nécessaires. Une sonde est transmise de domaine en domaine jusqu'à atteindre la destination (en respectant les contraintes) ou jusqu'à l'abandon de la recherche si un parcours d'une sonde devient trop long. Dans cet algorithme, les sondes ne peuvent pas être transmises à un nœud déjà visité. Lorsqu'une sonde arrive dans un nouveau nœud non-visité, ce dernier détermine les garanties de qualité de service qu'il peut fournir avec certitude pour cette requête. Les champs (chemin courant et poids du chemin courant) de la sonde sont mis à jour. Si le chemin courant ne satisfait pas les contraintes, les anciennes valeurs des champs sont restaurées et la sonde est renvoyée vers le dernier nœud visité. Si les contraintes sont toujours satisfaites par le chemin courant, alors la sonde est envoyée aléatoirement vers un nouveau domaine non-visité voisin du domaine courant. Enfin, si une sonde visite un nombre de nœuds supérieur à N_{hop}, la recherche est abandonnée. Nous avons aussi étudié une modification de l'algorithme présenté ici. Elle autorise la sonde à repasser plusieurs fois par un même nœud mais interdit de repasser plusieurs fois par le même lien.

Les résultats de la performance du réseau inter-domaine avec notre algorithme fournissant les routes multi-contraintes ont été confrontés à ceux obtenus avec les routes BGP standard donc sans possibilité d'y prendre en compte les contraintes [FAV08] (borne inférieure) et avec ceux obtenus par une recherche exhaustive de solutions parmi les routes (borne supérieure). En sachant que le problème de l'existence d'une route multi-contrainte peut être trivial lorsque les contraintes à satisfaire sont trop strictes ou trop relâchées [KM05] nous avons généré des requêtes ayant des contraintes très variées.

Les simulations ont été exécutées pour les topologies plates d'une part et hiérarchiques d'autre part. La génération de ces dernières a été assurée par notre outil SHIIP (cf. le paragraphe 2 du chapitre IV). Les routes à trouver concernent bande passante seule, délai seul, bande passante et délai, bande passante, délai et taux de pertes. Tous ces résultats ont été publiés dans [WT05, WT06a] et ils démontrent que notre algorithme fonctionne particulièrement bien dans les cas de des requêtes avec deux ou trois contraintes. Ils mettent aussi en évidence l'impact de routes de forme sans-vallée sur la performance du réseau bien que l'algorithme ne s'en serve pas. Cette influence se manifeste par le fait que l'algorithme fonctionnait parfois moins bien que BGP alors qu'il était toujours meilleur lorsque nous utilisions des routes quelconques. Nous pensons que cela est dû à la taille des routes qui sont plus longues dans les topologies hiérarchiques.

2 Multicast dans des réseaux optiques à circuits virtuels

Continuant des études dans le contexte de la QdS des applications massives distribuées dans le réseau CARRIOCAS (cf. le paragraphe 4.2 du chapitre II) nous nous sommes intéressés à la transmission *multicast* caractéristique des transmissions vidéo. Le protocole devant être déployé dans le réseau CARRIOCAS est GMPLS (*Generalized Multi-Protocol Label Switching*) [EM04]. Son fonctionnement est basé sur le principe de circuits virtuels.

2.1 Problème de routage

Le principe d'établissement de circuits virtuels rend l'utilisation d'une connexion pour chaque destination, comme décrit dans [SRV97, MYC+00], très inefficace dans le cas où des destinations

sont nombreuses [MZQ98]. Les schémas basés sur la construction d'un arbre de Steiner de poids minimal [Bea89, HR92] offrent des solutions plus efficaces mais ils supposent que tous les routeurs du réseau sont équipés de la fonctionnalité de dupliquer des paquets optiques. Cette fonctionnalité dans des routeurs de réseaux optiques nécessite la conversion O/E/O qui provoque l'augmentation du coût d'un tel routeur et de la consommation d'énergie et, potentiellement, l'introduction d'un délai supplémentaire. Pour ces raisons, des opérateurs de réseaux optiques veulent restreindre le nombre de routeurs capables de dupliquer les paquets optiques. Notre but est donc minimiser la bande passante utilisée par un arbre *multicast* avec une limite sur le nombre de routeurs pouvant dupliquer de paquets que nous appelons *nœuds diffusants* dans la suite de ce paragraphe. Dans [RTBW09] nous avons proposé de modéliser le réseau comme un graphe symétrique [Ber66] pondéré (le poids correspond à la bande passante disponible sur un lien) et nous avons formalisé le problème d'optimisation *diff_ tree* pour une demande de *multicast* demandant une unité de bande passante.

Problème III.4 Problème *diff_ tree*

Données : Un graphe orienté symétrique $G = (V, E)$ représentant un réseau, un ensemble des nœuds diffusants D_F, $D_F \subseteq V$ et une demande unitaire *multicast* $\varepsilon = (e, R)$, où e, $e \in V$, est la source de *multicast* et les éléments de R sont les destinataires de *multicast*.
Question : Existe-t-il un arbre enraciné en e dont la somme des arcs est minimale ?

Dans le même article nous avons démontré que ce problème d'optimisation est *NP*-complet. L'idée de cette preuve consiste à réduire le problème de couverture d'ensemble pondéré qui est NP-complet [Kar72] à notre problème.

La première question que nous nous sommes posée est la manière de placer de k nœuds diffusants dans le réseau. Ici nous avons opté pour une solution apportée par une des heuristiques proposées pour le problème de k-centres [MR05]. Le placement des nœuds diffusants sur les centres assure un placement homogène, c'est à dire minimisant la distance maximale entre un nœud diffusant (un centre) et tout autre nœud du réseau.

La seconde contribution consiste en une heuristique pour résoudre notre problème d'optimisation lié à *diff_ tree*. Il nous a semblé judicieux de la baser sur des algorithmes de flots. Pour ce faire, le graphe initial représentant le réseau (Figure III.1) est d'abord transformé en un graphe dit de diffusion (Figure III.2). Le graphe de diffusion est orienté, pondéré et il se compose de la source (l'émetteur du *multicast*), des nœuds diffusant, des nœuds destinataires du *multicast* et du puits. Le puits ne représente aucun nœud physique du réseau et il a été ajouté pour rendre le graphe "traitable" par des algorithmes de flots. Les arcs connectent la source à tous les nœuds diffusants. Tous les nœuds diffusants sont connectés entre eux dans les deux sens (formant une clique). Chaque nœud diffusant est lié avec les destinations du *multicast* et toute destination de *multicast* est connectée au puis.

La pondération des arcs est composée de deux valeurs et elle attribue à un arc :

- **le coût** égal à la longueur d'un chemin le plus court dans le graphe initial qui correspond au volume de trafic pouvant être acheminé entre des nœuds (les poids étiquetant des arcs entre des destinations et le puits "imaginaire" représentent le coût nul) et
- **la capacité** qui est introduite pour assurer que toutes les destinations du *multicast* soient atteintes ; elle est alors infinie sur tous les arcs sauf ceux qui connectent des destinations au puits et dont capacité est unitaire.

Cette pondération nous permet de capter le volume de bande passante disponible dans des segments de circuits virtuels entre des nœuds et assurer que chaque destination reçoit les données du *multicast* exactement une fois et une seule. Nous montrons qu'un arbre *multicast* enraciné dans la source permettant d'atteindre toutes les destinations et minimisant la bande passante utilisée sur l'ensemble des liens peut être construit avec une modification d'un des algorithmes classiques de recherche un

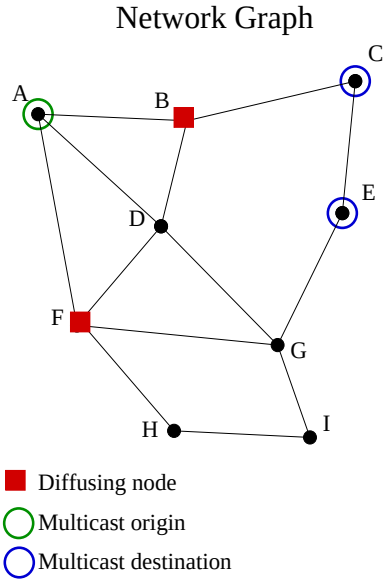

FIG. III.1: *Un réseau — graphe initial*

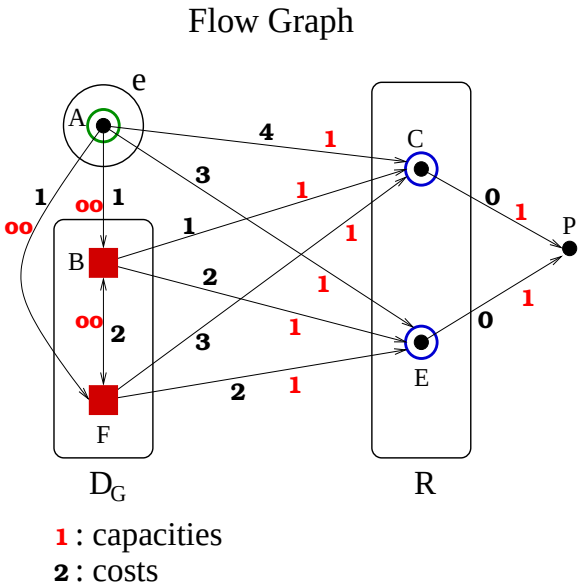

FIG. III.2: *Le graphe de flot correspondant au réseau initial*

flot de valeur maximale et de coût minimal comme celui proposé par Busacker et Gowen [Jun07]. Le but de cette modification est de ne pas utiliser plusieurs fois la même bande passante. Elle consiste à chaque étape à annuler le coût dans un arc qui vient d'être incorporé dans l'arbre qui est en train d'être construit. L'algorithme ne peut donc pas toujours donner précisément un arbre de coût minimal. Pour vérifier la performance de cet algorithme nous avons effectué des simulations intensives, dont la méthodologie sera brièvement discutée ci-dessous, qui montrent que notre heuristique est efficace. De plus sa complexité en $\mathcal{O}(n^4)$, où n est le nombre de routeurs d'un réseau, est raisonnable (la même complexité que celle de l'algorithme de Busacker et Gowen).

Afin d'étudier l'efficacité de notre heuristique nous avons opté pour la comparaison avec un algorithme exact pour des réseaux dans lesquels le nombre k de nœuds diffusants reste faible car sa complexité est pseudo-polynomiale en $\mathcal{O}(2^k)$. La méthodologie des expériences réalisées est décrite en détail dans [RTBW09]. Nous avons choisi un ensemble de *multicasts* de tailles différentes en variant la localisation des sources. Les expériences ont été faites pour des réseaux générés avec le générateur de topologies aléatoires BRITE [MLMB01] selon le modèle de génération de Waxman [Wax88] dont le nombre de nœuds variait de 100 à 500 et le nombre de nœuds diffusant changeait de six à douze. Nous avons observé que notre heuristique trouve toujours une solution et dans le pire des cas elle utilise 10% de bande passante de plus que la solution exacte. La dégradation de la performance de notre heuristique a été constatée pour des arbres consommant beaucoup de bande passante mais des destinations peu nombreuses. Ce phénomène peut être expliqué par le fait que notre heuristique n'utilise qu'une partie des nœuds diffusants qui sont à sa disposition. Le choix initial du sous-ensemble de nœuds diffusants sélectionné pour la construction d'un arbre peut être décisif pour l'efficacité de notre algorithme. Dans les situations dans lesquelles le nombre de destinations est important, notre algorithme utilise tous les nœuds diffusants et les poids des arbres qu'il trouve sont très proches de ceux obtenus par l'algorithme exact. Généralement nous avons constaté que notre algorithme est plus efficace quand le nombre de nœuds diffusants est petit par rapport au nombre de destinations d'un *multicast* car il utilisera tous les nœuds diffusants pour construire un arbre. Malgré ces différences, la performance de l'heuristique reste bonne dans toutes les expériences réalisées.

La problématique de transmissions *multicast* ouvre beaucoup de pistes de recherche intéressantes, notamment introduction de capacités des liens, d'une limite sur le nombre de nœuds diffusants autorisée sur une branche de l'arbre *multicast* où encore d'une limite sur le degré des nœuds diffusants.

2.2 Problème de dimensionnement

Le problème décrit dans le paragraphe précédent concernait le déploiement des arborescences *multicast* dans un réseau dont l'infrastructure a été établie au préalable. Or la connaissance des sources et destinations des transmissions *multicast* fréquentes et volumineuses peut inciter un opérateur d'un réseau à le configurer en adaptant le mieux à faire face à cette importante demande. Notre idée a été donc de trouver le moyen optimal pour déployer un nombre de nœuds diffusants fixe dans une arborescence de *multicast* donnée afin de minimiser la bande passante utilisée. Nous avons formulé le problème correspondant (Problème III.3) et nous l'avons résolu dans [RCT$^+$11b, RCT$^+$11a].

Nous avons donc formulé ce problème comme un problème d'optimalisation. L'idée pour le résoudre est venue grâce aux observations de certaines propriétés de la solution et leurs preuves [RCT$^+$11a].

Pour toute solution et pour tout nœud u nous avons défini un triplet, nommé *window* et composé : du nombre de chemins d'une solution arrivant à u, du nombre de nœuds diffusants, $|D^u|$, déployés dans le sous-arbre A^u et la quantité de la bande passante utilisée par A^u. Par exemple, *window* du nœud u dans Figure III.3 est égal à $(1, 2, 5)$. Nous avons démontré la formule pour calculer *window* d'un nœud à partir de *window* de son prédécesseur dans les cas où ce prédécesseur et un et un seul.

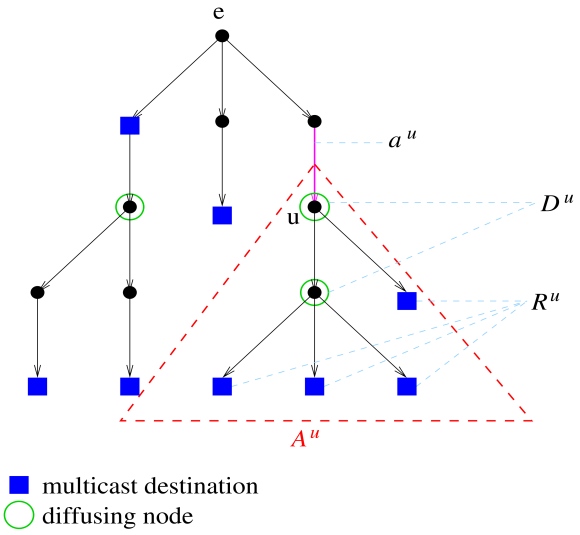

Fig. III.3: *Eléments d'une arborescence* multicast

Problème III.5 Problème *predefined_ tree*

Données : Un graphe orienté symétrique $G = (V, E)$ représentant un réseau, une demande unitaire *multicast* $\varepsilon = (e, R)$, où e, $e \in V$, est la source de *multicast* et les éléments de R sont les destinataires de *multicast*, une arborescence A_ε correspondant à cette demande *multicast* et un nombre entier positif k.

Question : Comment doit-on placer k nœuds diffusants dans A_ε pour que la quantité de la bande passante utilisée soit minimale ?

Nous avons également introduit l'ordre partiel relatif à un nœud donné v pour classer des solutions par rapport aux valeurs de leurs triplets *window* dans v. Cet ordre nous a permis par la suite de déterminer la qualité des solutions pour les nœuds dotés de plusieurs successeurs.

Ces observations nous ont mené à la proposition d'un algorithme exact basé sur l'approche dynamique. L'arborescence est parcourue en largeur mais un successeur, choisi arbitrairement, est traité différemment des autres, Figure III.4). Pour chaque nœud nous cherchons une solution qui est la meilleure selon l'ordre partiel que nous avons établi. La valeur de cette meilleure solution est gardée pour un nœud soit dans un vecteur $L(u)$ (dans les cas où u est considéré comme nœud diffusant) soit dans une matrice $M(u)$ (quand ce nœud est privé de fonctionnalité de duplication et diffusion de paquets). Cette distinction vient du fait que, de manière évidente, le nombre de chemins entrant dans un nœud diffusant est égal à un. La meilleure solution choisie pour un nœud donné sert à construire des solutions possibles pour ses prédécesseurs. Cette approche *bottom-up* trouve la solution optimale de l'arborescence entière comme une meilleure solution attribuée à sa racine.

La complexité en temps de notre algorithme optimal est en $\mathcal{O}(k^2|R|^2|V|)$ ce qui rend cette méthode, utilisée exclusivement sur l'étape de projet d'une infrastructure de réseau, très intéressant.

Bien que nous ayons prouvé l'optimalité de notre solution du problème III.5, le travail a été effectué pour estimer l'impact de la méthode de déploiement d'un arbre *multicast* sur l'efficacité du dimensionnement d'un réseau. Parmi des méthodes heuristiques communément approuvées pour résoudre ce problème NP-complet nous avons choisi deux : celle qui établit des chemins les plus courtes entre la source d'une demande *multicast* et toutes ses destinations (ShP) et celle qui résout le problème classique d'arbre de Steiner dans des graphes non-orientés (StT). Nous avons implémenté

Procedure `Mat_Vec_Filling`
1. **If** u is a leaf **then** attribute the "unitary" $M(u)$ and $L(u)$ to u
2. **else**
3. choose arbitrarily v which is one of the successors of u in A^u ;
4. `First_Succ_Mat_Vec(u,v)` ;
5. mark v ;
6. **While** there is a successor of u in A^u which has not be marked yet **do**
7. choose arbitrarily w among the non-marked successors of u in A^u ;
8. `Others_Succ_Mat_Vec(u,w)` ;
9. mark w
10. **endWhile**
11. **endIf**

Fig. III.4: *Procédure du parcours en largeur* `Mat_Vec_Filling`

l'algorithme 2-approché proposé dans [TM80] pour StT.

Les expériences ont été réalisées pour des réseaux de 200 nœuds générés avec BRITE [MLMB01] selon le modèle de Waxman [Wax88]. Le poids moyen des arbres *multicast* a été estime avec précision 5% et niveau de confiance $\alpha = 0.05$. Le choix d'une source de *multicast* et, par la suite, des destination a été effectué suivant une distribution uniforme.

Les résultats complets sont présentés dans [RCT+11a]. Ici nous monteront uniquement leur sélection. Dans Figures III.5 et III.6 nous observons l'impact de la présence de quatre nœuds diffusants sur le poids moyen des *multicasts* en fonction du nombre de destinations. La première observation intéressante que nous pouvons déjà faire est que la méthode ShP, trouvant des chemins les plus courtes, produit des arbres plus légers que la méthode StT. Le gain obtenu pour la construction des arborescences selon ShP (Figure III.5) est significatif (approximativement 31% pour 32 destinations). Il est même plus important dans le cas de la construction basée sur le problème de Steiner, Figure III.6 (approximativement 65% pour 16 destinations). Nous pouvons donc constater que l'introduction des nœuds diffusants nivelle l'avantage initial de la méthode ShP.

Après avoir fixé le nombre de destinations à 20 nous avons poursuivi nos expériences afin d'estimer la minimisation de la bande passante utilisée, obtenue par l'ajout de nœuds diffusants (Figures III.7 et III.8). Bien que les poids des arborescences construites selon la méthode ShP soient presque deux fois inférieurs à ceux générés par la méthode StT, l'introduction de nœuds diffusants de nouveau "gomme" ce désavantage. Par exemple, le déploiement de 15 nœuds diffusants réduit le poids moyen des arbres ShT de 20% et le même nombre de nœuds diffusant dans un réseau avec des *multicast* construits selon la méthode StT voit la réduction du poids moyen de 75%. Nous constatons que les arborescences StT peuvent devenir plus intéressantes que ShP de point de vue de l'utilisation de la bande passante grâce à l'introduction de nœuds diffusants.

Nous avons également établi le critère permettant de prendre la décision du choix de la méthode de la construction des arborescences. Pour les réseaux et les *multicasts* dont les paramètres, nombre de nœuds diffusants et nombre de destinations, se trouvent au-dessus de la ligne droite (Figure III.9) nous conseillons la méthode ShP.

Nous aimerons poursuivre notre recherche dans ce contexte en implémentant l'algorithme pour trouver la solution de l'arbre Steiner plus précis que celui que nous avons utilisé (par exemple

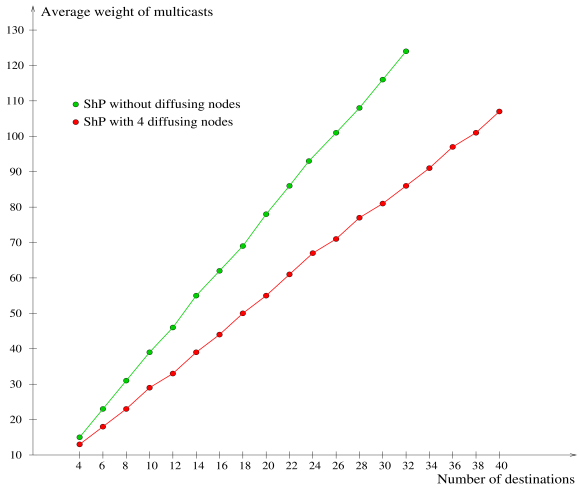

Fig. III.5: *Poids moyen de A_ϵ^{ShP} en fonction du nombre de destinations avec quatre nœuds diffisants et sans nœuds diffusants*

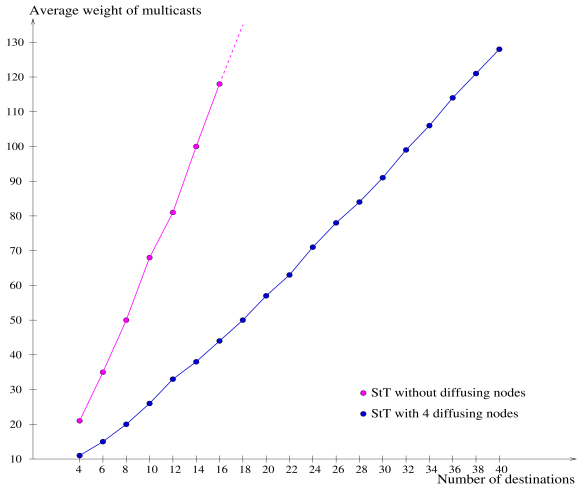

Fig. III.6: *Poids moyen de A_ϵ^{StT} en fonction du nombre de destinations avec quatre nœuds diffisants et sans nœuds diffusants*

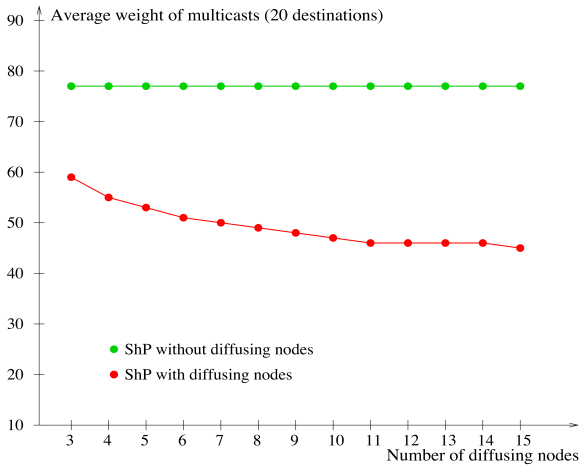

Fig. III.7: *Poids moyen de A_ϵ^{ShP} en fonction du nombre de nœuds diffusants et sans nœuds diffusants pour* 20 *destinations*

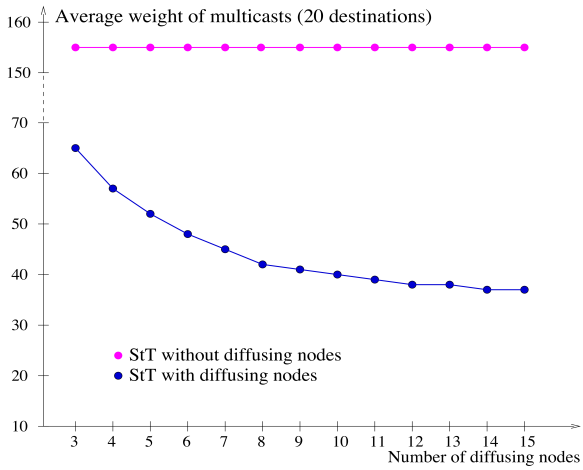

Fig. III.8: *Poids moyen de A_ϵ^{StT} en fonction du nombre de nœuds diffusants et sans nœuds diffusants pour* 20 *destinations*

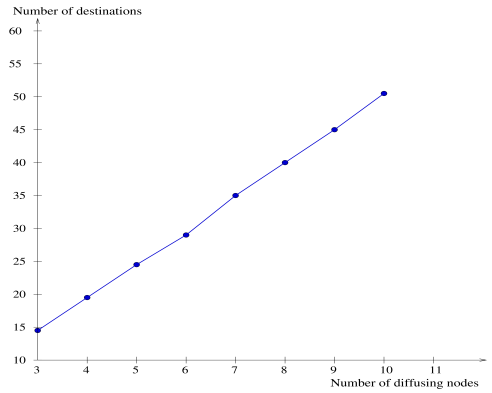

Fig. III.9: *Ligne critique déterminant l'utilisation des méthodes ShP et StT de la construction des arborescences (ShP au-dessus cette ligne, StT au-dessous cette ligne)*

[RZ00, BGRS10]). Un autre axe potentiel de recherche consiste à explorer de réseaux dont le *tree-width* [KT06] est borné. Il serait intéressant de vérifier si la solution du problème III.5 dans le *treewidth* borné donnait un placement de nœuds diffusants au préalable pour le problème III.4 plus efficace qu'une solution du problème de k-centres utilisée pour l'instant (cf. le paragraphe 2.1 de ce chapitre).

3 Dimensionnement d'un anneau tout optique

Nous avons repris la problématique concernant le dimensionnement des réseaux dans le cadre du projet DOROthéE (DimensiOnnement des Réseaux métropolitains tout-Optiques à faible consommation Energétique) financé par le biais de Digiteo-DIM. Nous nous sommes penchés sur la minimisation du coût d'infrastructure CAPEX (*CAPital EXpenditure*) d'un réseau métropolitain WDM/TDM basé sur une nouvelle architecture de nœuds proposée par Alcatel-Lucent [CNSA10]. Cette architecture, POADM (*Packet Optical Add-Drop Multiplexer*) dont le schéma est illustré par Figure III.10 prévoit "le contournement" d'un nœud par certaines longueurs d'onde. Cette propriété est introduite par la séparation des composants responsables de la lecture Rx (*receiver*) et des composants responsables de l'écriture Tx (*transmitter*). Le découplage Rx/Tx constitue une différence majeure par rapport aux multiplexeurs OADM classiques (cf. le paragraphe 3 du chapitre II). Par conséquent, bien que l'injection des données puisse se faire sur toutes les longueurs d'onde grâce à un système de lasers (*Tunable Lasers*, TL dans Figure III.10), les Rxs peuvent être attachés uniquement à un sous-ensemble de longueurs d'onde. La connaissance préalable de la nature du trafic prévu transporté par le réseau permet donc de réduire le nombre de Rxs nécessaires et de baisser le CAPEX.

Dans un premier temps nous avons considéré le problème de minimisation du nombre de longueurs d'onde prenant comme contrainte le nombre de Rxs nécessaires. Nous pouvons fixer cette contrainte à son minimum car le calcul du nombre minimal de Rxs est direct comme ceci est présenté dans la définition formelle du problème (Problème III.6) que nous citons ici après notre article [PTWB11].

contrainte de flot : Pour tout couple de nœuds (i, j) :

$$\forall i, j \in V, \sum_{k \in \lambda} T_k[i, j] = T[i, j].$$

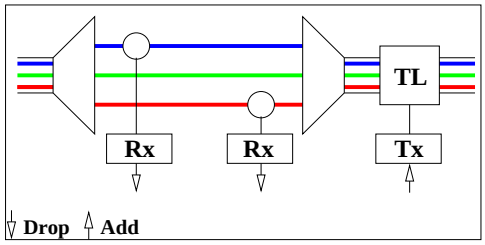

Fig. III.10: *Architecture POADM*

Problème III.6 Problème MWLP (*Minimum WaveLength Problem*)

Données : Un circuit élémentaire [Ber66] $G = (V, E)$, une matrice de trafic T avec $T[i, j]$ indiquant la quantité de trafic envoyé du nœud i au nœud j, un ensemble $\lambda = \{\lambda_1, \lambda_2, ..., \lambda_K\}$ de longueurs d'onde, $K \in \mathbb{N}$ et la capacité d'une longueur d'onde C, $C \in \mathbb{N}$
Nous utilisons le terme *affectation* pour décrire une opération consistant à décomposer la matrice de trafic T en un ensemble de matrices T_k de la même dimension que T et à associer T_k à une longueur d'onde k.
Question : Existe-il une affectation du trafic donné par la matrice T aux longueurs d'onde de l'ensemble λ telle qu'elle respecte trois contraintes (flot, capacité, Rxs) détaillées ci-dessous ?

contrainte de capacité : Après avoir noté $load_k(x)$ la quantité de trafic transportée par un arc x sur une longueur d'onde k :

$$\forall k \in \lambda, \forall x \in E, \text{load}_k(x) = \sum_{i,j \ s.t. \ x \in path(i,j) \in G} T_k[i, j] \leq C.$$

contrainte de Rxs : Nous notons ici w_i^- l'ensemble des longueurs d'onde sur lesquelles le nœud i doit lire pour pouvoir recevoir tout le trafic qui lui est destiné. Nous observons que cette valeur donne directement le nombre minimal des Rxs indispensable à la lecture de tout trafic destiné au nœud i :

$$\forall i \in V, w_i^- = \{k \in \lambda | \forall s \in V, T_k[s, i] \neq 0\},$$

$$\forall i \in V, |w_i^-| = \left\lceil \frac{\sum_j T[j, i]}{C} \right\rceil.$$

Dans [PTWB11] nous avons prouvé la NP-complétude de ce problème en utilisant pour la réduction polynomiale le problème *Bin-Packing* [MT90]. Nous y avons aussi proposé une heuristique pour résoudre le problème III.6. Son idée consiste à imaginer chaque longueur d'onde dans un anneau à n nœuds comme un cube dans n dimensions dont la longueur des arêtes est égale à C. Nous supposons que les nœuds de cet anneau sont numérotés de 1 à n et une i-ème dimension est associée avec un i-ème arc (celui entre les nœuds i et $i+1$). Nous formons des objets à empaqueter dans ces n-cubes à partir de transmissions qui passent entre deux nœuds. Un objet regroupe les transmissions ayant la même destination, ordonnées des requêtes les plus longues aux requêtes plus courtes. Le *packing* se fait avec le découpage horizontal (toutes les "tranches", sauf potentiellement la dernière, ont une hauteur C) mais il n'admet ni translation ni rotation.

La validation de cet algorithme a été effectuée pour des réseaux de tailles variées et avec différentes distributions pour déterminer la quantité de trafic et les couples (source, destination). Des résultats complets présentés dans [PTWB11] nous avons choisi ceux qui montrent le nombre de longueurs

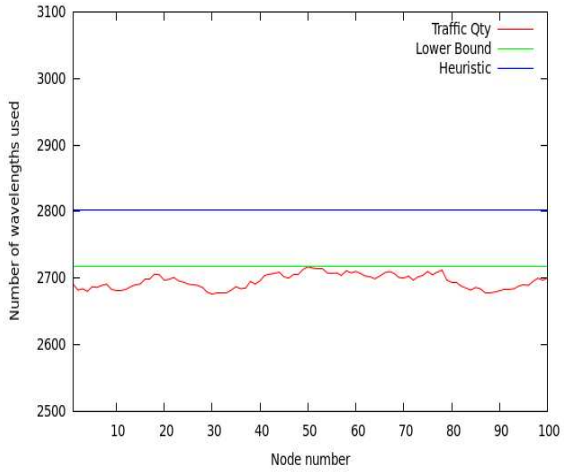

Fig. III.11: *Résultat de notre heuristique pour distribution du volume de trafic uniforme et distribution de sources/destinations uniforme*

d'onde obtenus dans un anneau à 100 nœuds par rapport à la borne inférieure prise comme la quantité de trafic sur l'arc le plus contraint. Les matrices de trafic ont été construites avec une distribution uniforme du volume de trafic et les deux types de distribution spatiale :

1. uniforme pour le sources, uniforme pour les destinations,
2. uniforme pour les sources et RGR (*Rich Get Richer*) [BA99] (cf. le paragraphe 2 du chapitre IV) pour les destinations.

Dans le cas de distribution spatiale uniforme (Figure III.11) notre algorithme produit les résultats à peine 3% moins bon que la borne inférieure. Son comportement est même meilleur lorsque la distribution spatiale du trafic suit une loi RGR (Figure III.12) : 1.7% de dégradation seulement par rapport à la borne inférieure.

La suite de ces travaux est en cours et consiste à résoudre le problème de minimisation du nombre de Rxs.

4 Qualité de service dans des réseaux ad hoc de la sécurité civile

Dans le cadre du projet RAF (Réseaux Ad hoc à Forte efficacité) nous avons abordé la problématique de la qualité de service dans des réseaux ad hoc devant être déployés pour assurer la communication entre des acteurs de sécurité publique intervenant sur les scènes de catastrophes. Un moyen d'assurer la QdS demandée consiste à utiliser un mode de multiplexage comme TDMA ou FDMA. Les exigences principales de la part des utilisateurs d'un réseau ad hoc pour la sécurité civile sont les garanties de 1) connectivité et 2) disponibilité de ressources. En même temps des réservations doivent être faites en tenant compte des interférences. Contrairement à la majorité des articles consacrés à ce domaine, nous avons décidé d'aborder l'assurance de la QdS en deux étapes séparées :

1. la détection et la résolution des interférences et
2. l'allocation de ressources au sens strict du terme.

Nous nous sommes inspirés ici des articles [BBC+05, BCG+05] qui ont proposé une séparation similaire dans le contexte d'un système à mémoire partagée.

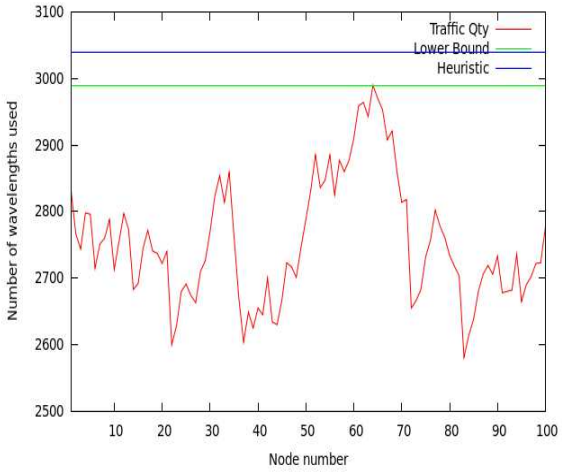

Fig. III.12: *Résultat de notre heuristique pour distribution du volume de trafic uniforme, distribution sources uniforme et distribution de destinations RGR*

4.1 Détection et correction des interférences à deux sauts

Nous avons utilisé l'approche visant l'allocation de ressources sur des liens (et non sur des nœuds) car selon [Grö00] la réservation sur des liens garantie une meilleure réutilisation spatiale. Dans notre modèle les liens sont orientés afin de représenter mieux des communications asymétriques.

Notre réseau ad hoc est donc modélisé par un graphe orienté $G = (V, E)$ appelé *le graphe de connectivité* dont l'arc $(t, r) \in E$ représente la connexion entre l'émetteur t et le récepteur r. Les nœuds du graphe G ont des canaux attribués (soit des *time-slots*, soit des fréquences) et nous supposons que ces ressources sont indépendantes. Dans le contexte de TDMA cette indépendance de ressources est assurée par une synchronisation forte. Dans le contexte de FDMA elle est réalisée par l'attribution de fréquences orthogonales. Nous supposons aussi que le graphe G est connexe.

Dans la plupart des articles traitant le problème d'allocation dans des réseaux sans fil *slotés* modélisant le réseau par le graphe de connectivité (par exemple [STT02, KUHN03, HT04, RWMX06, SBMIY07]), la zone d'interférences est égale à la zone de transmission. Néanmoins, vu la nature de l'atténuation de signal radio, ce signal, déjà trop faible en dehors de la zone de transmission, peut être toujours suffisamment fort pour interférer avec un autre signal. L'hypothèse que la zone d'interférences est plus large que la zone de transmission a été posée notamment dans [WWL$^+$06, KN05, KMPS05, DV07].

La modélisation des interférences est faite par un *graphe de conflits* $G_C = (V_C, E_C)$ pour lequel il existe une relation entre V_C et E du graphe de connectivité G. Un arc $(e_1, e_2) \in E_C$ existe seulement si un émetteur d'e_i et un récepteur d'e_j, $i, j \in \{1, 2\}, i \neq j$ sont à distance de deux sauts.

Comme nous l'avons signalé ci-dessus, nous utilisons le principe du système à mémoire partagée et sans échanges directs de messages entre des nœuds. Nous attribuons une variable d'état par canal pour chaque nœud. Un nœud connaît ses variables et sait lire celles de ses voisins à un saut. Il peut changer les valeurs de ses propres variables uniquement. Un nœud exécute un programme constitué de variables et d'actions gardées. L'état d'un nœud est défini par les valeurs de ses variables. Une configuration du système à un moment donné est constituée des états de tous les nœuds. L'exécution du système est une séquence de configurations. Le passage d'une configuration à une autre est réalisé par l'exécution simultanée des actions gardées dans les nœuds dans lesquels les gardes sont satisfaites.

Tab. III.1: *Description de la variable* channelState

Etat	Portée	Signification
T	locale	un *slot* utilisé localement pour la transmission
R	locale	un *slot* utilisé localement pour la réception
N_R	un saut	au moins un voisin utilise le *slot* pour la réception
N_{T1}	un saut	au moins un voisin utilise le *slot* pour la transmission
N_{TR}	un saut	un *slot* utilisé par au moins un voisin-émetteur et par exactement un voisin-récepteur
N_{TR+}	un saut	un *slot* utilisé au moins un voisin-émetteur et plus que deux voisins-récepteurs
N_{T2}	deux sauts	un *slot* utilisé par un émetteur à distance de deux sauts
F	—	aucun émetteur à distance de deux sauts et aucun récepteur dans le voisinage à un saut

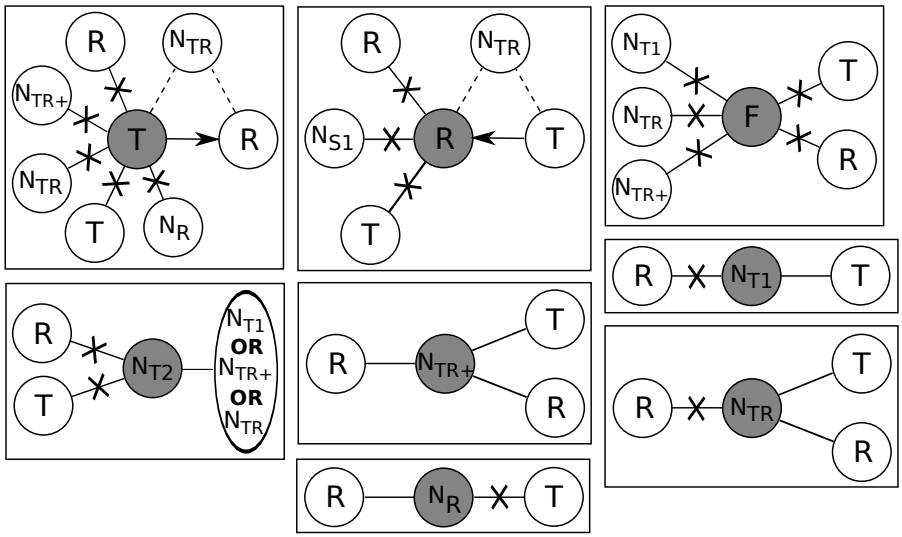

Fig. III.13: *Relations entre des nœuds voisins*

Nous reprenons d'après [BDPV07, DIM97] la notion de *ronde*. Une ronde est définie comme un ensemble minimal de configurations dans lesquelles tout nœud est activable (autrement dit : tout nœud a au moins une action dont la garde est satisfaite) au moins une fois. La variable d'état d'un nœud pour un canal est appelée *channelState*. Dans le Table III.1 nous présentons sa description.

La Figure III.13 schématise les relations entre les nœuds d'un voisinage. Dans cette figure les lignes continues indiquent les conditions qui doivent être satisfaites, les lignes barrées indiquent les conditions qui sont interdites et les lignes pointillées indiquent les conditions dont la satisfaction n'est pas nécessaire mais qui n'introduisent pas de conflits.

Dans l'article [PTBV10] nous avons proposé un algorithme de détection et de correction des conflits d'interférences à deux sauts. Afin d'interdire l'activation de plusieurs actions changeant différemment la valeur de la variable *channelState* nous proposons des prédicats introduisant une priorité des actions. L'information concernant des états d'un voisinage compris dans $\{N_R, N_{TR}, N_{TR+}, N_{T1}\}$ est plus prioritaire que celle concernant des états dans $\{N_{T2}, F\}$. La priorité suit la relation $N_{TR+} > N_{TR} > N_{T1} \geq N_R$ et $N_{T2} > F$. Dans notre article nous avons démontré la sûreté et la vivacité de l'algorithme proposé. Notre algorithme garantit que dans l'état stationnaire un système ne contient

pas de conflits et cet état stationnaire est atteint au pire après les cinq rondes.

Selon notre intuition notre algorithme fonctionne bien dans des cas de mobilité des agents et est toujours capable de résoudre des conflits dans un voisinage à deux sauts.

4.2 Allocation de ressources avec interférences à deux sauts

La suite des travaux présentés dans le paragraphe ci-dessus a concerné l'allocation des ressources dans un réseau "nettoyé" d'interférences à deux sauts. Etant placés dans le contexte de réseaux ad hoc gérés par le protocole OFDMA nous prenons comme une unité d'allocation un *Ressource Block*, (RB), vu comme couple (temps, fréquence). Notre but est de fournir un RB par arc pour assurer la connectivité exigée par la sécurité civile. Les travaux résumés dans ce paragraphe sont décrit en détail dans [PTBV11].

Pour modéliser un réseau nous avons choisi, comme précédemment, *un graphe de connectivité*, un graphe orienté $G = (V, E)$ qui permet représenter l'allocation sur les liens en sens opposés. Notre but a été de maximiser la réutilisation spatiale des ressources. Pour obtenir la réutilisation spatiale maximale on aurait pu être tenté par l'utilisation d'un RB donné tous les trois sauts. Néanmoins, le fait que toute nouvelle allocation doit tenir compte de celles qui existent déjà, rend le temps de convergence d'un tel algorithme inacceptablement long. Pour cette raison nous avons opté pour certains compromis. Ils ont été réalisés par l'introductions des priorités exprimées par des poids attribués à des nœuds.

Le poids $W_n = dg^-(n) + \sum_{v \in \text{voisins}(n)} dg^-(v)$ d'un nœud n, où $dg^-(n)$ indique le nombre d'arcs sortants de n et voisins(n) indique l'ensemble des nœuds pour lesquels n est prédécesseur ou successeur, donne le nombre de liens potentiellement affectés par une allocation d'un RB sur un lien sortant du nœud n.

La prise en compte des allocations déjà établies est réalise par un autre poids, $W'_n = W_n - nombre\ de\ liens\ sortants\ de\ n\ sur\ lesquels\ des\ allocation\ sont\ déjà\ faites.$

Nous avons supposons également que chaque nœud est identifiable par son numéro ID.

Notre algorithme prend la décision si une allocation peut être faite pour un nœud n voulant émettre selon le principe suivant :

1. si $W'(n) < W'_{n'}$ pour tout $n' \in$ voisins(n), alors l'allocation sur un lien sortant de n est accordée,
2. s'il existe n', $n' \in$ voisins(n) tel que $W'_{n'} < W'_n$, alors n doit attendre son tour d'allocation,
3. si n et ses voisins ont W' identiques, la procédure est réitérée selon les poids W et, dans le cas où elle échoue, la décision est prise arbitrairement selon le numéro d'ID de nœud.

Cet algorithme fonctionne conjointement avec l'algorithme de résolution de conflits d'interférences à deux sauts qui a été enrichi par rapport à sa version décrite dans le paragraphe 4.1 de ce chapitre par l'ajout de quatre variables d'état permettant de tenir compte de la priorité d'allocations effectuées. Les modifications de l'algorithme de résolution de conflits d'interférences originel sont précisées dans [PTBV11].

Ci-dessous nous présentons les résultats de simulations obtenus pour une topologie aléatoire (un tirage uniforme pour deux coordonnées). L'espace de simulation est 1000×1000. La portée radio varie afin de conserver le degré moyen des nœuds constant malgré des topologies de taille différente selon la formule $\frac{\text{dg\_moyen} \cdot \text{surface\_de\_simulation}}{\pi(N-1)}$, où N est le nombre de nœuds du réseau. Un nœud déclare la volonté d'émettre toutes les 500 ms. Une ronde est un intervalle de temps nécessaire pour que tous les nœuds puissent envoyer leur message. Les moyennes du temps de convergence et du nombre de RBs alloués sont calculées à partir des échantillons de taille 1000 (Figures III.14 et III.15,

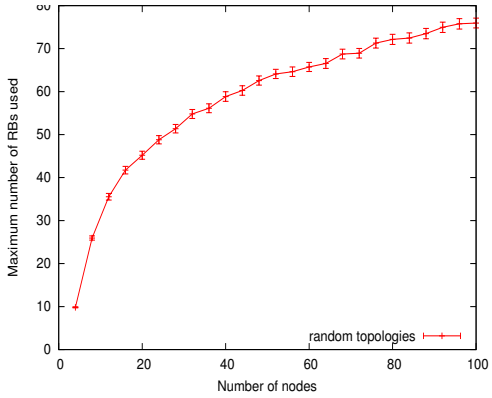

Fig. III.14: *Temps de convergence moyen en rondes pour un réseau de topologie aléatoire et ses intervalles de confiance*

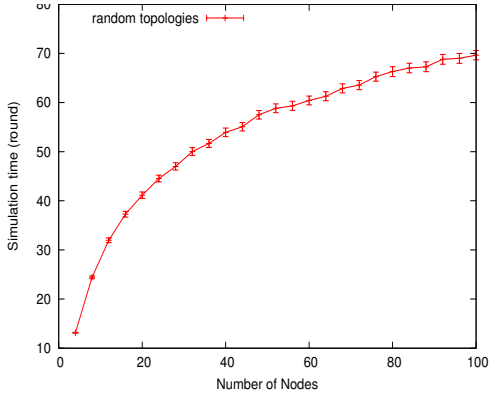

Fig. III.15: *Nombre de RBs en moyenne pour un réseau de topologie aléatoire et ses intervalles de confiance*

intervalles de confiance 95%). Ces résultat montrent que notre algorithme utilise la même ressource plusieus fois et leur quantité allouée a tendence à se stabiliser bien que la taille du réseau croisse.

Nous voulons continuer l'étude de l'allocation dans le contexte de la QoS plus large, en tenant compte de la matrice de trafic.

Chapitre IV

Descriptions de modèles

La validation de solutions algorithmiques proposées afin de d'introduire la gestion de la qualité de service dans des réseaux nécessite l'évaluation des performances. Cette évaluation est faite à partir des modèles analytiques produisant des résultats exacts ou empiriques dont les mesures sont estimées à partir des résultats collectés suite à des réalisations de processus stochastiques par un simulateur à événements discrets. Nous avons travaillé sur des modèles markoviens de très grande taille ainsi que sur des modèles topologiques et de mobilité permettant d'incorporer dans la description du modèle fourni à l'entrée du simulateur la spécificité du système étudié.

Le paragraphe 1 de ce chapitre présente l'approche décompositionnelle appliquée à la modélisation markovienne. Le deuxième paragraphe décrit l'idée d'introduction de la hiérarchie induite par les relations commerciales existantes dans le réseau inter-domaine et le logiciel SHIIP (*Supélec Hierarchy Inter-domain Inducting Program*) ainsi que sa version étendue aSHIIP (*autonomous Supélec Hierarchy Inter-domain Inducting Program*). Le dernier paragraphe est consacré à la présentation du modèle de mobilité de secouristes proposé dans le cadre du projet RAF.

1 Formalismes compositionnels

L'évaluation de performance de systèmes se base sur le traitement numérique de leurs modèles, il existe donc un vaste domaine consacré à la description algébrique de modèles. D'un côté un formalisme descriptif doit être suffisamment expressif pour pouvoir incorporer dans un modèle des propriétés d'un système modélisé. De l'autre côté il doit permettre la transformation d'un modèle en forme adaptée à la solution numérique et le calcul de paramètres de performance du système analysé.

Nous nous sommes principalement intéressés aux formalismes algébriques SAN (*Stochastic Automata Network*) [Pla84] (nous avons consacré le travail à ce formalisme dans notre thèse [Tom98]) et PEPA (*Performance Evaluation Process Algebra*) [Hil96], les deux applicables à des modèles markoviens et également à la description formelle du processus de simulation exprimée par GSMP (*Generalised Semi-Markov Process*) [Mat62].

Le formalisme de réseau d'automates stochastiques SAN a été introduit par Brigitte Plateau dans sa thèse [Pla84]. La théorie a été décrite formellement dans [Ati92] et les applications dédiées notamment pour les réseaux de télécommunications ont été présentées dans [Que94]. La méthode SAN est principalement adaptée aux modèles markoviens en temps continu et dans ce manuscrit nous nous limitons à cette échelle du temps.

L'idée novatrice du formalisme SAN consiste à représenter une chaîne de Markov multidimensionnelle par un ensemble de processus stochastiques de dimension plus petite, en pratique — unidi-

mensionnels. Ces processus sont décrits par des automates stochastiques, leur ensemble est nommé d'un réseau d'automates stochastiques et ils peuvent être vus comme les descriptions du comportement des composants de la chaîne de Markov globale, un automate décrivant un composant de l'état. Les transitions exponentielles pouvant se produire dans des automates stochastiques et formant la description d'une chaîne de Markov globale sont classées en deux groupes par rapport à leur portée. Les transitions intervenant uniquement dans un automate sans influencer les autres sont appelées les transitions locales. Afin de pouvoir changer d'état dans plusieurs automates en même temps, les transitions globales sont utilisées. Les transitions globales interviennent donc dans un sous-ensemble d'automates stochastiques formant un réseau et elles les synchronisent. Les transitions globales sont déclenchées par les événements synchronisants qui sont identifiés par leurs noms. La théorie de SAN permet d'utiliser les taux de transitions fonctionnels, c'est à dire que une valeur courante d'un taux de transition, soit locale, soit globale, peut dépendre des états d'un sous-ensemble d'automates d'un réseau. Tout automate stochastique est donc défini par un ensemble de matrices de transitions responsables de changements de son état (une matrice pour des transitions locales et une matrice pour chaque événement synchronisant appliqué à cet automate). La composition de matrices associées aux automates d'un réseau se fait avec les moyens offerts par l'algèbre tensorielle (l'algèbre de Kronecker) [Dav81]. Le générateur markovien est calculé par la somme tensorielle de matrices de transitions locales additionnée au produit tensoriel de matrices de synchronisation normalisées au préalable.

La dimension d'une matrice de chaîne de Markov globale est ici un facteur décisif déterminant le temps de génération de ses éléments qui devient non-négligeable. Certains auteurs [PF91] optent pour la génération *on-the-fly* (après avoir permuté les lignes et colonnes du générateur pour faciliter les calculs) : un réseau d'automates stochastiques existe toujours uniquement dans sa forme "dispersée", ses éléments sont calcules au besoin, au fur et à mesure des multiplication vecteur-matrice consécutives effectuées lors du processus de solution itératif. Dans nos travaux nous avons choisi une autre approche consistant à générer une matrice, la stocker selon le schéma compact (index d'une colonne, nombre d'éléments non-nuls dans cette colonne n, (index de ligne d'un élément non-nul, valeur de cet élément)$_{1,2,\ldots,n}$) et trouver sa solution stationnaire comme vecteur propre gauche correspondant à la valeur propre 1 par la méthode Arnoldi [Arn51, Saa80] basée sur la projection de la matrice originale dans l'espace de Krylov [Saa81, Saa91, Saa95] dont la dimension est plus petite que celle de l'espace initial. Notre méthode de solution nous apparaît plus efficace que une approche basée principalement sur des synchronisations car nous utilisons dans nos modèles SAN les transitions fonctionnelles qui seront discutées ci-dessous. La préférence vers l'utilisation fréquente des transitions fonctionnelles a été également la raison pour ne pas utiliser le logiciel PEPS [BBF$^+$03] pour résoudre nos modèles.

Comme exemple du modèle contenant des transition fonctionnelles nous présentons ici un modèle d'un simple mécanisme d'accès illustré dans Figure IV.1. Il gère deux classes de clients dont la première est prioritaire. Les paquets de la première classe sont générés par une source exponentielle dont le taux d'emission est égal à λ_1. Les paquets de la deuxième classe arrivent aussi selon une distribution de Poisson, mais avec taux de λ_2. Les temps de service pour les deux classes sont distribués exponentiellement et égaux à μ_1 et μ_2, respectivement. Les paquets qui ne sont pas privilégiés doivent attendre des périodes d'absence de paquets prioritaires pour pouvoir passer. Ils disposent d'un tampon dont la taille est fixée à b.

Le réseau d'automates stochastiques décrivant ce modèle est présenté dans Figure IV.2. L'automate SA$^{(0)}$ est responsable de la présence de paquets de première classe. L'automate SA$^{(1)}$ compte de paquets de deuxième classe ($b = 2$ dans cet exemple). Le symbole χ indique la fonction caractéristique dont la valeur est 1 si son argument est vrai et 0 dans le cas contraire. Notre fonction χ prend comme argument une expression logique dépendante de l'état de l'automate SA$^{(0)}$, $s^{(0)}$.

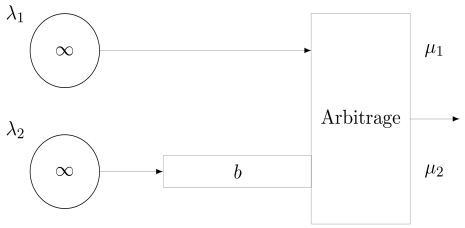

Fig. IV.1: *Schéma d'un mécanisme d'accès simple avec deux classes de paquets*

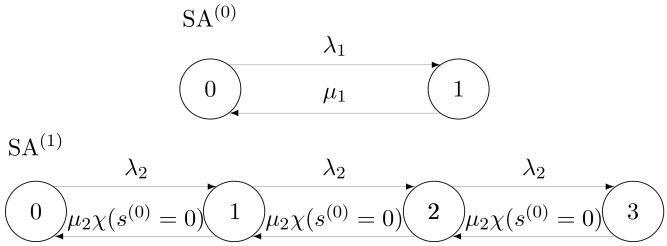

Fig. IV.2: *Automate $SA^{(0)}$ décrivant le processus représentant le nombre de paquets de première classe et automate $SA^{(1)}$ décrivant le processus représentant le nombre de paquets de deuxième classe, $b = 2$*

Cette fonction bloque la possibilité de passage pour des paquets de la deuxième classe quand un paquet privilégié est présent. Nous observons que le système est décrit sans utilisation de transitions synchronisantes.

Les générateurs locaux de ces deux automates sont donnés par les matrices :

$$\boldsymbol{Q}^{(0)} = \begin{bmatrix} -\lambda_1 & \lambda_1 \\ \mu_1 & -\mu_1 \end{bmatrix} \qquad \text{et}$$

$$\boldsymbol{Q}^{(1)} = \begin{bmatrix} -\lambda_2 & \lambda_2 & 0 & 0 \\ \mu_2\chi(s^{(0)}=0) & -(\mu_2\chi(s^{(0)}=0)+\lambda_2) & \lambda_2 & 0 \\ 0 & \mu_2\chi(s^{(1)}=0) & -(\mu_2\chi(s^{(0)}=0)+\lambda_2) & \lambda_2 \\ 0 & 0 & \mu_2\chi(s^{(0)}=0) & -\mu_2\chi(s^{(0)}=0) \end{bmatrix}.$$

Les valeurs des éléments fonctionnels du générateur local $\boldsymbol{Q}^{(1)}$ seront connues au moment du choix de leur place dans le générateur global suivant la définition des opérations tensorielles généralisées. Dans le cas de notre exemple le générateur global sera :

$$\boldsymbol{Q} = \begin{bmatrix} \Sigma_0 & \lambda_2 & 0 & 0 & \lambda_1 & 0 & 0 & 0 \\ \mu_2 & \Sigma_1 & \lambda_2 & 0 & 0 & \lambda_1 & 0 & 0 \\ 0 & \mu_2 & \Sigma_1 & \lambda_2 & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & \mu_2 & \Sigma_2 & 0 & 0 & 0 & \lambda_1 \\ \mu_1 & 0 & 0 & 0 & \Sigma_3 & \lambda_2 & 0 & 0 \\ 0 & \mu_1 & 0 & 0 & 0 & \Sigma_3 & \lambda_2 & 0 \\ 0 & 0 & \mu_1 & 0 & 0 & 0 & \Sigma_3 & \lambda_2 \\ 0 & 0 & 0 & \mu_1 & 0 & 0 & 0 & -\mu_1 \end{bmatrix},$$

où $\Sigma_0 = -(\lambda_1 + \lambda_2)$, $\Sigma_1 = -(\lambda_1 + \lambda_2 + \mu_2)$, $\Sigma_2 = -(\lambda_1 + \mu_2)$, $\Sigma_3 = -(\lambda_2 + \mu_1)$.

La possibilité d'utiliser les transition fonctionnelles dont le taux effectif est donné par une fonction a été étudiée en détail dans notre thèse [Tom98] dans laquelle nous avons proposé les principes de leur application dans des modèles concernant des réseaux de télécommunications. Nous avons ensuite utilisé cette approche dans certaines études de cas présentées dans le chapitre V en nous basant notamment sur les résultats de [QT99] permettant la modélisation efficace des systèmes en termes de SAN.

Il n'existe pas une seule et unique représentation d'une chaîne de Markov multidimensionnelle comme réseau d'automates stochastiques. De notre point de vue, la représentation la plus avantageuse mène à la génération de la matrice markovienne en utilisant des opérations les moins coûteuses au sens du temps d'exécution. Dans [QT99] nous avons notamment présenté l'algorithme permettant la conversion des transitions locales fonctionnelles en transactions globales dont les taux sont constants et nous avons illustré son fonctionnement en proposant un modèle de file d'attente multiple dotée d'un mécanisme *Round Robin* [KR03, Sta04]. On peut remarquer qu'un réseau d'automates stochastiques sans dépendances fonctionnelles nécessite plus de multiplications pour trouver des éléments du produit tensoriel mais d'un autre côté l'effort de calcul des taux de transitions fonctionnelles dépend de leur complexité et du type de données utilisées puisque la taille de la représentation numérique des données influence fortement le temps de traitement. Pour un utilisateur, il est plus pratique de décrire son modèle avec des transitions fonctionnelles qui sont conceptuellement plus naturelles et dans le cas de leur inadaptabilité numérique, de les transformer automatiquement en transitions synchronisées de taux constants.

Les modèles markoviens des systèmes informatiques doivent inclure énormément de paramètres qui introduisent des dépendances entre des composants d'un système réel. Cela mène à des modèles dont les matrices sont de taille trop grandes pour pouvoir les traiter numériquement dans leur totalité. Les contraintes limitant l'utilité de modèles markoviens proviennent de deux critères : la mémoire, puisque le nombre d'éléments non-nuls devient trop important et le temps, puisque le temps de construction d'une matrice de générateur ainsi que le temps nécessaire pour qu'une méthode itérative converge deviennent inacceptablement longs. Les méthodes compositionnelles telles comme PEPA [Hil96] peuvent apporter une solution au problème d'explosion de nombre d'états d'une chaîne de Markov modélisant un système complexe. Cette solution consiste à traiter un modèle comme une composition de sous-modèles qui peuvent être traités indépendamment.

Comme exemple nous donnons ici un modèle du système constitué d'un processus et d'une ressource. Pour commencer nous décrivons un processus qui effectue deux activités en alternance, l'utilisation d'une ressource *util* et l'exécution de calculs *tâche*, dont les durées respectives sont décrites par les distributions exponentielles avec les paramètres r_1 et r_2.

$$\text{Processus} \stackrel{def}{=} (util, r_1).(tâche, r_2).\text{Processus}.$$

Le composant de la ressource effectue deux activités en alternance. La première est liée à son utilisation par un processus (le même type d'action *util*), la deuxième est une activité interne de la ressource provoquée par sa mise à jour, *mise-à-jour*.

$$\text{Ressource} \stackrel{def}{=} (util, r_3).(mise\text{-}à\text{-}jour, r_4).\text{Ressource}$$

Grâce à l'opération de collaboration nous pouvons construire notre modèle comme suit (l'exécution simultanée des activités dont le type d'action est *util*) :

$$\text{Système} \stackrel{def}{=} \text{Processus} \underset{\{util\}}{\bowtie} \text{Ressource}.$$

Nous pouvons observer la facilité de construction d'autres modèles à partir des briques disponibles, par exemple en utilisant une compoition parallèle :

$$\text{UnAutreSystème} \stackrel{def}{=} (\text{Processus}||\text{Processus}) \underset{\{util\}}{\bowtie} \text{Ressource}.$$

Le modèle dont les composants sont statistiquement indépendants et dont la solution décomposée dans l'état stationnaire est la même que celle de la chaîne globale est appelé modèle avec la solution en forme produit (*product form solution*) [vD93]. Dans [TH00b] nous avons analysé la possibilité de transformer des modèles formalisées en PEPA afin d'obtenir les bornes de leurs facteurs de performance en forme produit. La possibilité d'amalgamer certaines transitions dans des modèles exprimés en PEPA ont été étudiés dans [TH00a, HT00]. Les agregations peuvent être effectuées uniquement dans de certains cas qui ne dénaturent pas la propriété markovienne. L'existence de mémoire se traduit par le fait que la fonction de hasard attribuée à une transition n'est pas exponentielle. Ces résultats ont encouragé le travail concernant la formalisation de processus sous-adjacents d'un processus de simulation, un processus semi-markovien généralisé, *Generalised Semi-Markov Process, GSMP* [Tom02].

La méthode de Réseaux Stochastiques avec de transitions fonctionnelles nous a permis de construire des modèles de Markov de système complexes (voir chapitre V). Ces experiences montrent que cette méthode est un moyen efficace de construire la matrice de générateur des chaînes de Markov de très grande taille.

2 Induction de la hiérarchie du réseau inter-domaine dans une topologie plate

Comme cela a été annoncé dans le paragraphe 1 du chapitre II, les relations commerciales existantes entre des domaines influencent la diffusion de messages d'annonces de routes du protocole BGP (routes *sans-vallée* [Gao01]) et imposent une hiérarchie entre les domaines du réseau [BEH+07, GFJG01, SARK02]. Les analyses de tables de routage de BGP menées par les chercheurs cités ont mis en évidence l'existence de la stratification du réseau Internet en cinq couches. La couche la plus haute, numéro 0, nommée aussi le cœur, est composée exclusivement des domaines qui ne sont clients d'aucun autre domaine et ils peuvent uniquement être interconnectés par de liens *P2P*. En général, les domaines d'une couche i, $i = 1, 2, 3, 4$ sont clients de domaines d'une couche supérieure $i - 1$.

En travaillant sur l'introduction des aspects de la qualité de service dans le réseau inter-domaine Internet (cf. le paragraphe 1 du chapitre III) nous avons ressenti le besoin d'avoir des jeux de topologies inter-domaines avec hiérarchie de taille inférieure de celle d'Internet car la validation des algorithmes proposés aurait été limitée à une seule configuration trop coûteuse en termes de temps de simulation. N'ayant pas à notre disposition de générateur de topologies inter-domaines avec hiérarchie nous avons décidé d'en écrire un. Nous nous somme partis du principe d'introduction de la hiérarchie dans une topologie inter-domaine plate générée selon un des algorithmes connus.

Nous avons exclu au départ les algorithmes de génération structuraux basés sur la représentation de propriétés globales d'un réseau tels que Tiers [Doa96] et Transit-Stub [ZCB96] car le nombre de couches dans des topologies qu'ils produisent est fixé à trois et deux, respectivement. Nous nous sommes focalisés sur des approches de génération basées sur le degré comme Inet [JCJ00], les modèles d'Aiello et al. [ACL00], de Barabási et Albert (BA) [BA99], le modèle étendu de Barabási et Albert (eBA) [AB00] et le modèle généralisé de préférence linéaire (GLP) [BT02]. L'objectif de chacun de ces modèles est d'être représentatif d'Internet et particulièrement de la loi de puissance qui caractérise les degrés des domaines. En effet [TRGW02] présente la comparaison des topologies

produites à partir des générateurs de graphes structuraux et de graphes basés sur le degré. Selon cette étude les propriétés globales (particulièrement la longueur caractéristique des chemins) sont mieux représentées dans les topologies générées à partir des modèles de graphes basés sur le degré.

Pour générer les topologies plates nous avons d'abord repris un outil couramment utilisé par la communauté scientifique BRITE [MLMB01] qui offre la génération de topologies selon des algorithmes mentionnés ci-dessus (BA, eBA et GLP) ainsi que l'algorithme de génération de Waxman [Wax88].

Les critères que nous avons sélectionnés comme propriétés de la hiérarchie souhaitables sont les suivants :

- la couche la plus haute, le cœur, est une clique uniquement composée de liens *P2P*,
- plus une couche est haute dans la hiérarchie, plus le nombre de nœuds qui la constitue est réduit et plus leur degré moyen est grand,
- les domaines d'une couche (à l'exception de ceux du cœur) sont reliés à la couche supérieure par au moins un lien *C2P*.

L'induction de la hiérarchie dans une topologie plate est réalisée en cinq étapes :

1. génération d'une topologie plate,
2. construction d'un cœur,
3. construction des différentes couches,
4. optimisation de la distribution des nœuds dans les couches,
5. déduction du type des relations.

Le point 1 a été initialement réalisé avec le logiciel BRITE en utilisant quatre modèles de génération mentionnés ci-dessus. Pour le point 2 nous tenons compte du fait que le cœur d'Internet est un graphe très dense [TDG+01, SARK02] et le cœur construit est la plus grande clique composée des nœuds ayant les plus grands degrés. L'étape 3 est particulièrement difficile à réaliser à cause du manque de résultats de mesures différents de ceux obtenus pour la totalité des domaines d'Internet et à cause du fait que le problème similaire, enrichi par la connaissance de tailles de couches est déjà *NP*-complet. La preuve de *NP*-complétude du problème de construction de couches avec leurs tailles données consistait à montrer une équivalence avec le problème de la couverture de sommets [Kar72, Gib85] est présentée en détails dans [TW10b]. La structure sans-vallée des routes contenues dans Internet implique que tout nœud est relié à un sommet du cœur par une succession de liens de *C2P*. Nous supposons que chaque nœud est relié au cœur par un plus court chemin. Nous utilisons cette propriété et cette hypothèse pour construire une heuristique pour choisir la couche à laquelle appartient chacun des nœuds. Cette heuristique ne tient pas compte du degré d'un nœud attribué à une couche et par conséquent, des couches hautes en hiérarchie ont tendance à être "trop grandes". A la lumière de [SARK02, CGJ+04] nous voyons que plus une couche est proche du cœur, plus sa taille est petite et plus le degré moyen des nœuds la constituant est grand. Dès l'étape 4 nous essayons de "repousser" les nœuds dont le degré est plus faible vers les couches basses. L'algorithme glouton proposé, décrit en détails dans [WT07a], garantit qu'un nœud changeant de couche conserve un lien avec sa couche d'origine et tous les clients de ce nœuds gardent au moins un fournisseur appartenant à la couche d'origine. La dernière étape est faite selon deux règles simples :

- une arête joignant deux nœuds d'une même couche est transformée en relation *P2P*,
- une arête joignant deux nœuds de couches différentes est transformée en relation *C2P*, le fournisseur étant le nœud de la plus haute couche.

Nous avons réalisé l'introduction de la hiérarchie selon notre algorithme dans des séries de topologies plates des tailles variant de 40 à 2000 nœuds générées avec quatre (Waxman, BA, eBA, GLP) schémas disponibles dans BRITE. Les mesures prises sur des hiérarchies obtenues avec l'intervalle de précision de 5% avec un niveau de confiance de 95% ont concerné :

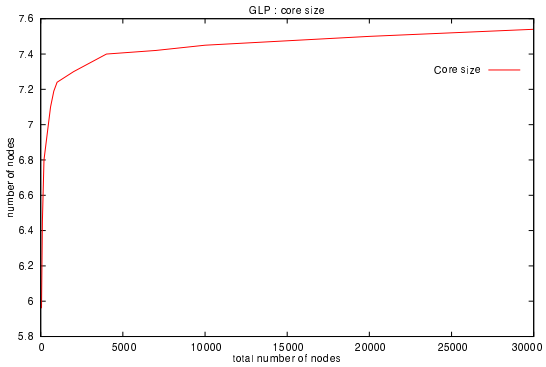

Fig. IV.3: *Taille du cœur dans des topologies générées par GLP avec le degré moyen de domaines de 4.4 en fonction de la taille de topologie.*

- la taille du cœur,
- le nombre moyen de nœuds par couche,
- le degré moyen des nœuds dans une couche,
- le nombre de relations *P2P* et *C2P*,
- la longueur des routes.

L'analyse de résultats [WT06b, WT07a, Wei07] mène à la conclusion que notre algorithme introduit une hiérarchie satisfaisant les critères dans les topologies de toutes les tailles testées générées selon le modèle eBA et il donne aussi des hiérarchies désirées dans des réseaux d'au moins 400 domaines générés selon le schéma GLP.

Le logiciel SHIIP (*Supélec Hierarchy Inter-domain Inducting Program*) étant l'implémentation de l'algorithme présenté était disponible sur Internet sous licence libre. Il pouvait donc bénéficier à la communauté scientifique qui nécessite un tel générateur [FPSS02, BEHV05]. Le générateur a été utilisé dans les projets RNRT Actrice et ACI SR2I [BMM09, BBM09].

Suite à des perspectives de la large utilisation du générateur SHIIP et au fait que le générateur BRITE de topologies plates n'est plus maintenu par ses auteurs, nous avons conçu la nouvelle version de SHIIP, aSHIIP, où "a" signifie "autonomous" [TW10a, TW10b]. aSHIIP génère des topologies plates selon les mêmes modèles que BRITE et il est doté en plus du modèle d'Aiello [ACL00]. Par rapport à la version précédente, aSHIIP cherche le cœur qui est "presque" une clique (un graphe auquel manquent au plus deux arêtes pour être une clique). aSHIIP sait traiter des topologies de grande taille, comparable avec celle d'Internet qui a actuellement, selon les mesures prises par Caida [cai], plus de 33 mille domaines. La topologie captée par Caida [cai] est caractérisée par le degré moyen de nœuds égal à 4.4, le degré maximal égal à 2502 et le coefficient γ de la loi de puissance décrivant la distribution de degré de nœuds égal à $2,331$. Cette valeur de γ a été estimée de la propriété théorique de la loi de puissance $k_{\max}^{\mathrm{PL}} = n^{\frac{1}{\gamma-1}}$ [DM03]. Le modèle GLP avec son paramètre de probabilité $p = 0.55$ génère typiquement des topologies avec le degré moyen de 4.4, le degré maximal qui s'approche de mille et le coefficient γ qui est égal à 2.54. Son paramètre β responsable du choix du nœud d'attachement pour un lien à attacher, n'a d'influence, ni sur le degré moyen, ni sur γ. Le paramètre β a, néanmoins, un impact sur le degré maximal d'une topologie. Par conséquent, ce paramètre a une influence sur la position du cœur et la position des couches. Nous

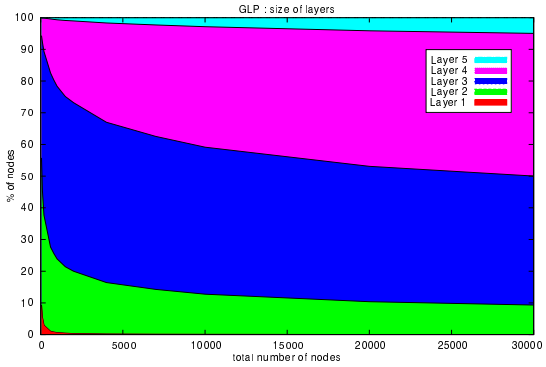

Fig. IV.4: *Distribution de nœuds sur les couches dans des topologies générées par GLP avec le degré moyen de domaines de 4.4 en fonction de la taille de topologie.*

conseillons de fixer $\beta = 0.75$. A notre avis, le modèle GLP est le mieux adapté pour la génération des topologies représentatives pour Internet.

Dans Figures IV.3, IV.4 et IV.5 nous montrons la taille du cœur, la distribution de nœuds sur les couches et le degré moyen de nœuds par dans des topologies générées par GLP avec le degré moyen de domaines de 4.4 en fonction de la taille de topologie. Le pourcentage de relation *P2P* dans des topologies générées par aSHIIP se trouve entre 16% et 17%. Ce résultat est conforme avec celui présenté dans [DKF+07] qui est de 16.1%.

L'algorithme d'introduction de la hiérarchie dans la topologie Caida [cai] la découpe en six couches de taille : 0.04% (14 domaines, le cœur), 9.16%, 63.69%, 25.28%, 1.79% et 0.03%. Le degré moyen sur les couches commençant par le cœur est 780.69, 20.9, 2.6, 1.95, 1.67 et 1.18. Selon nous deux dernières petites couches devraient être absorbées par la couche *Tier-4* dont les nœuds ont un degré comparable avec ceux des couches résiduelles.

Le logiciel aSHIIP est disponible sur la page `http://wwwdi.supelec.fr/software/ashiip/`.

Nous prévoyons des axes importants du développement d'aSHIIP. Dans un premier temps nous voulons incorporer dans l'éventail des méthodes de générations de topologies plates celle proposée récemment par des membres du consortium Caida [SFK+10].

Nous pensons que notre générateur devrait être enrichi de paramètres décrivant des temps de passage par les domaines. Ces temps font partie de l'évaluation de la satisfaction de la QdS. Une autre raison pour laquelle l'étude des durées de traversée est nécessaire est le fait qu'Internet a subi depuis récemment de changements architecturaux non négligeables. Ces changements sont provoqués par la stratégie commerciale des fournisseurs de contenu qui déploient leurs propres infrastructures WAN [GALM08] afin d'amener leurs réseaux au plus proche des clients. Par conséquent les phénomènes observés sont 1) un degré important de domaines de contenu [OPW+08] et 2) des contournements des domaines du cœur par nombreuses routes [GALM08]. Il n'est pas exclu que l'intensification de ses tendances doive impliquer la révision de la structure hiérarchique. Nous somme convaincus que l'analyse du temps de passage par les domaines peut apporter les éléments permettant de détecter des déformations de la hiérarchie.

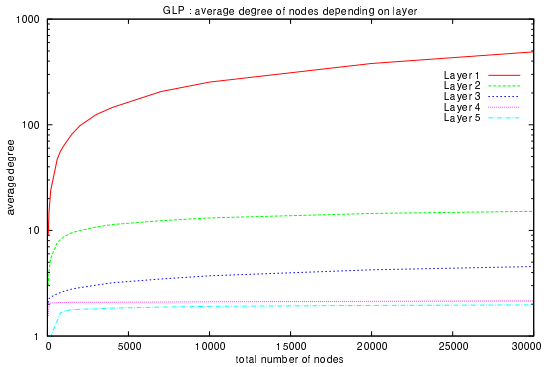

Fig. IV.5: *Degré moyen de nœuds par dans des topologies générées par GLP avec le degré moyen de domaines de 4.4 en fonction de la taille de topologie.*

3 Modèle de mobilité pour des réseaux ad hoc de la sécurité civile

Lors de notre travail concernant le projet RAF qui vise la garantie de la qualité de service dans des réseaux ad hoc de la sécurité civile, nous étions confrontés au problème de modélisation de déplacements de secouristes sur un lieu d'intervention. Les secouristes se déplacent à pieds et travaillent en équipes (pompiers, agents de police, personnel médical, etc.). Leurs interventions ont souvent lieu dans un milieu dans lequel existent de nombreux obstacles (bâtiments, décombres, etc.) qu'ils doivent contourner.

Les contraintes énumérées ci-dessus nous ont convaincus que le modèle de mobilité *Random Way Point* (RWP) [JM96] qui, selon [KCC05] est le plus répandu dans les simulations notamment celles faites avec ns-2, n'est pas adapté aux exigences du projet RAF. Le manque de réalisme de ce modèle a été constaté dans de multiples études, par exemple [GSB+08, BH04, Bet01, RMSM01, AGPG+07]. Nous avons trouvé que le modèle qui reflète le mieux les déplacements humains avait été proposé dans [RSH+08]. Ce modèle utilise des distributions de Levy pour décrire des longueurs de déplacement (*flight lenght*) et des durées de pauses (*pause time*) des agents mobiles. Les distributions de Levy sont en pratique approximées par des lois de puissance qui sont observées dans des activités humaines (voir paragraphe 2 de ce chapitre).

Les auteurs de [HGPC99] ont proposé un modèle de mobilité de groupes (*Reference Point Group Mobility*, RPGM). Bien que les déplacements des agents mobiles soient basés sur RWP, ce modèle suppose que les agents mobiles sont divisés en groupes dont chacun a un chef (*leader*). Le chef d'un groupe peut être soit un agent particulier soit un point virtuel. Dans le contexte de la sécurité civile la présence d'un chef de groupe désigné est fortement justifiée.

L'influence des obstacles sur des trajectoires de secouristes a été étudiée dans [AGPG+07] par le biais du graphe de visibilité [PS85]. Nous avons constaté que cette solution n'empêche pas des secouristes modélisés de s'approcher trop près des obstacles qui peuvent être potentiellement dangereux (par exemple, la présence du feu ou l'instabilité des murs). La solution qui nous est apparue la plus adaptée pour permettre à des secouristes de contourner des obstacles est basée sur le diagramme de Voronoï et a été décrite dans [JBRAS03]. Un diagramme de Voronoï sert à constituer "un réseau de pistes" à emprunter par des secouristes pour manœuvrer entre des obstacles. Les coins des obstacles sont pris comme points de référence pour construire ce diagramme. Néanmoins, nous

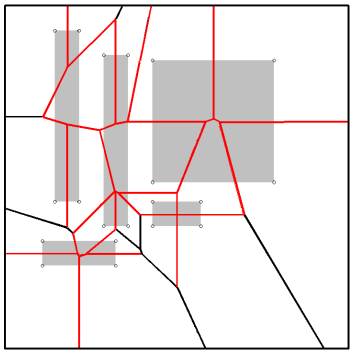

Fig. IV.6: *Un exemple de carte avec un diagramme de Voronoï selon l'approche classique (points de référence — coins des obstacles)*

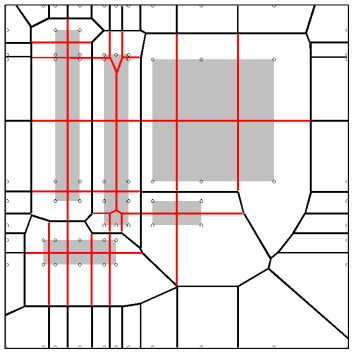

Fig. IV.7: *Un exemple de carte avec un diagramme de Voronoï selon notre approche (points de référence — coins des obstacles et leurs projections sur les obstacles voisins et le bord)*

avons constaté que cette solution n'est pas satisfaisante au bord de la surface de simulation et pour certaines configurations d'obstacles (par exemple de longs obstacles parallèles et séparés par une courte distance).

Pour satisfaire les exigences posées à des simulations dans le projet RAF nous avons proposé dans [PTV10, PTV11] un modèle de mobilité compositionnel. Nous avons repris le modèle de mobilité de groupes RPGM en remplaçant la distribution des mouvements RWP par des distributions de Levy. Pour obtenir "un réseau de pistes" plus réaliste que celui offert par le diagramme de Voronoï classique, nous en avons proposé une modification. En plus de l'utilisaton des coins des obstacles comme points de référence nous prenons aussi des projections des coins sur les obstacles voisins et sur le bord de la surface de simulation. Figures IV.6 et IV.7 illustrent sur un exemple les différences entre deux approches.

Le modèle que nous avons proposé a satisfait des contraintes imposées dans des scénarii du projet concernant des interventions de la sécurité civile. Il est utilisé dans des simulations dans l'industrie.

Chapitre V

Etudes de cas

Dans ce chapitre vous présentons les modèles markoviens proposés pour des éléments de réseaux (cf. les paragraphes 1, 2 et 3 du chapitre II). Les modèles markoviens d'un commutateur d'un réseau local sont présentés dans le paragraphe suivant. Le paragraphe 2 décrit de *switches* tout-optiques. Le paragraphe 3 concerne la gestion d'accès à des réseaux optiques.

1 Commutateur multiservice dans un réseau local

Nous avons analysé un commutateur de réseau de l'architecture STAR de l'Intranet corporatif [Tél96]. Nous avons distingué quatre classes de paquets : signalisation, voix, vidéo et données non-prioritaires. La qualité de service à assurer pour ces classes est différente car les paquets de signalisation sont sensibles à la fois aux pertes et au délai, les paquets de voix et vidéo sont principalement sensibles au délai et les paquets de données n'ayant pas de contraintes peuvent passer par le réseau en fonction de la disponibilité de la bande-passante (meilleur effort). Afin de pouvoir traiter des paquets dans un commutateur d'une manière conforme à la qualité de service demandée, nous avons proposé un mécanisme de file prioritaire avec un serveur (Figure II.1 dans chapitre II). Pour représenter les demandes différentes de qualité de service, des paquets de quatre classes sont stockés dans un commutateur dans trois tampons correspondant chacun au niveau de la qualité de service. Les paquets de voix et vidéo partagent le même tampon. Le serveur traite des paquets moins prioritaires uniquement à condition que des paquets plus prioritaires soient absents. Nous supposons que le temps nécessaire pour changer l'attachement du serveur à un tampon est négligeable.

Le modèle markovien de ce commutateur a été réalisé avec la méthode des Réseaux d'Automates Stochastiques (cf. le paragraphe 1 du chapitre IV) en profitant de la particularité de pouvoir utiliser des taux de transitions fonctionnels qui avaient été étudiés en détail dans [Tom98]. Afin de disposer de flots de paquets dont l'intensité varie en temps, nous avons représenté leurs sources par des groupes de sources ON/OFF identiques et indépendantes qui sont modélisés par des MMPPs.

Dans le premier modèle construit, le temps de passage des paquets par le commutateur est distribué exponentiellement. Grâce à la modularité de l'approche de la méthode de Réseaux d'Automates Stochastiques, nous avons pu ensuite enrichir ce modèle en traitement de durées constantes en y introduisant l'approximation par une distribution d'Erlang.

La structure d'état de la chaîne de Markov du premier modèle est la suivante : $\vec{s} = \left(s^{(0)}, s^{(1)}, s^{(2)}, s^{(3)}, s^{(4)}, s^{(5)}, s^{(6)}, s^{(7)}\right)$ où quatre premiers composants $s^{(k-1)}$ indiquent le nombre de paquets de quatre classes k, $k = 1, 2, 3, 4$ et quatre derniers $s^{(k+3)}$ décrivent le nombre de sources ON/OFF de chaque classe actives (l'état de la chaîne modulante pour l'émission de paquets de classe k, $k = 1, 2, 3, 4$). Le comportement d'un composant $s^{(i)}$, $i = 0, 1, \ldots, 7$ est défini par un automate $\mathrm{SA}^{(i)}$.

Les automates décrivent l'état des tampons dans le commutateur sont représentés graphiquement sur Figures V.1 ($SA^{(0)}$ et $SA^{(3)}$) et V.2 ($SA^{(1)}$ et $SA^{(2)}$). Les paquets de vidéo et de voix partagent le même tampon de capacité $B_{2,3}$. Il faut donc respecter la condition $s^{(1)} + s^{(2)} \leq B_{2,3}$. Figure V.3 illustre les automates responsables de la gestion de nombre de sources ON/OFF actives pour chaque classe de paquets.

Le temps de traitement dans le commutateur "pur" d'un paquet de classe k, $k = 1, 2, 3, 4$, est distribué exponentiellement avec la moyenne $1/\mu^{(k-1)}$. Les taux de transition responsables du traitement et de l'envoi de paquets de classe k par le commutateur doivent tenir compte de l'état de tampons, autrement dit de l'état de la chaine globale $\vec{s}$, afin de pouvoir réaliser la politique prioritaire qui y est implémentée. Ces taux fonctionnels seront notés $\mu^{(k-1)}(\vec{s})$. Des paquets de classes 2 et 3 (vidéo et voix) sont stockés dans le tampon commun. L'ordre des paquets n'est pas mémorisé pour ne pas faire exploser le nombre d'états de la chaîne de Markov globale. Pour cette raison $\mu^{(1)}(\vec{s})$ et $\mu^{(2)}(\vec{s})$ contiennent la probabilité qu'un paquet d'une classe donnée soit le premier à partir. Des paquets de la classe la plus prioritaire pouvant toujours partir, donc $\mu^{(0)}(\vec{s}) = \mu^{(0)}$, les paquets de vidéo et voix peuvent partir uniquement quand les paquets de la classe la plus prioritaire sont absents, $\mu^{(j)}(\vec{s}) = \frac{s^{(j)}}{s^{(1)}+s^{(2)}}\mu^{(j)}\chi(s^{(0)} = 0)$, $j = 1, 2$, et les paquets de données peuvent être uniquement émis quand il n'y a pas de paquets d'autres classes en attente, $\mu^{(3)}(\vec{s}) = \mu^{(3)}\chi(s^{(0)} = 0)\chi(s^{(1)} = 0)\chi(s^{(2)} = 0)$. Le symbole χ indique ici la fonction caractéristique dont la définition est rappelée dans le paragraphe 1 du chapitre IV.

Les taux de génération exponentielle de paquets de classe k, $k = 1, 2, 3, 4$ par une seule source ON/OFF sont notés λ_k. Le taux de génération de paquets de classe k par un groupe de N_{k+1} sources ON/OFF, défini par $\lambda^{(k-1)}(\vec{s})$, dépend du nombre de sources étant dans la période d'émission déterminé par l'automate $SA^{(k+3)}$ représentant la modulation markovienne de l'activité de ce groupe. Les taux $\lambda^{(1)}(\vec{s})$ et $\lambda^{(2)}(\vec{s})$ tiennent également compte du fait que des paquets vidéo et voix partagent le même espace de stockage. Sous les conditions posées ci-dessus nous écrivons $\lambda^{(k-1)}(\vec{s}) = \lambda_k s^{(k+3)}$, $k = 1$ et $k = 4$, et $\lambda^{(k-1)}(\vec{s}) = \lambda_k s^{(k+3)}\chi(s^{(1)} + s^{(2)} \leq B_{2,3})$ pour $k = 2$ et $k = 3$.

Le modèle qui vient d'être présenté ne contient pas d'événements synchronisants ayant les dépendances entre les automates réalisées uniquement par le biais de transitions fonctionnelles.

L'introduction d'une distribution d'Erlang modélisant le temps de traitement constant nécessite la présence d'événements synchronisants. L'état de la nouvelle chaîne contient onze éléments, les composants $s^{(8)}$, $s^{(9)}$ et $s^{(10)}$ indiquant une phase courante de service d'Erlang respectivement pour la signalisation, les paquets vidéo et voix confondus et des données. Les automates $SA^{(j)}$, $j = 8, 9, 10$ (Figure V.4) gèrent les distributions d'Erlang dont le nombre de phases est R_ν, $\nu = 1, \{2, 3\}, 4$ et le taux d'une phase est $\mu_R^{(j)}$, $j = 8, 9, 10$. Un événement synchronisant e_j, $j = 0, 1, 2$ est déclenché au moment où un service dont la durée est définie par une distribution d'Erlang se termine et un paquet quitte le commutateur. L'événement e_0 correspond au départ d'un paquet de la classe la plus prioritaire ($SA^{(0)}$ et $SA^{(8)}$ réagissent ensemble), l'événement e_2 correspond au départ d'un paquet de la classe la moins prioritaire ($SA^{(3)}$ et $SA^{(10)}$ sont synchronisés), et l'événement e_1 est responsable de l'envoi soit d'un paquet vidéo, soit d'un paquet de voix. Dans ce dernier cas nous observons soit la synchronisation entre $SA^{(1)}$ et $SA^{(9)}$, soit entre $SA^{(2)}$ et $SA^{(9)}$. La description détaillée des transitions fonctionnelles du modèle avec l'approximation des distributions déterministes est donnée par Eq.(V.1)–(V.8).

$$\begin{array}{rcl} \lambda^{(0)}(s) & = & \lambda_1 \cdot s^{(4)} \\ \mu^{(0)}(s) & = & \mu^{(0)} \\ p^{(0)}(s) & = & 1 \cdot \chi(s^{(8)} = R_1) \end{array} \qquad \text{(V.1)}$$

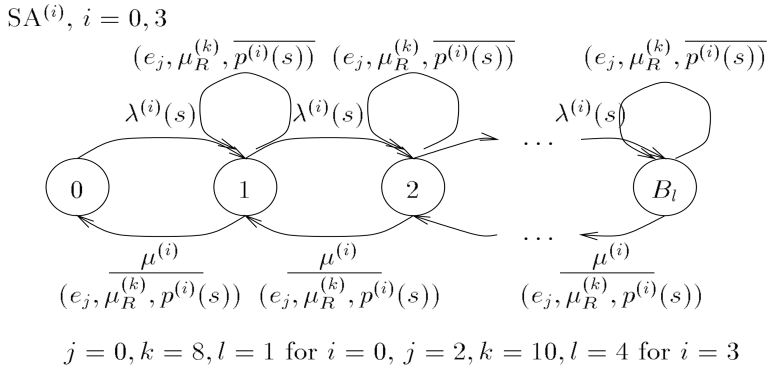

Fig. V.1: *Automate $SA^{(0)}$ représentant le nombre de paquets de signalisation et automate $SA^{(3)}$ représentant le nombre de paquets de données, Eq. (V.1), (V.4)*

$$\begin{aligned}
\lambda^{(1)}(s) &= \lambda_2 \cdot s^{(5)} \chi(s^{(1)} + s^{(2)} < B_{2,3}) \\
\mu^{(1)}(s) &= \tfrac{s^{(1)}}{s^{(1)}+s^{(2)}} \mu^{(1)} \chi(s^{(0)} = 0) \\
p^{(1)}(s) &= \tfrac{s^{(1)}}{s^{(1)}+s^{(2)}} \chi(s^{(0)} = 0) \chi(s^{(9)} = R_{2,3}) \chi(s^{(2)}_{\text{succ}} = s^{(2)})
\end{aligned} \tag{V.2}$$

$$\begin{aligned}
\lambda^{(2)}(s) &= \lambda_3 \cdot s^{(6)} \chi(s^{(1)} + s^{(2)} < B_{2,3}) \\
\mu^{(2)}(s) &= \tfrac{s^{(2)}}{s^{(1)}+s^{(2)}} \mu^{(2)} \chi(s^{(0)} = 0) \\
p^{(2)}(s) &= \tfrac{s^{(2)}}{s^{(1)}+s^{(2)}} \chi(s^{(0)} = 0) \chi(s^{(9)} = R_{2,3}) \chi(s^{(1)}_{\text{succ}} = s^{(1)})
\end{aligned} \tag{V.3}$$

$$\begin{aligned}
\lambda^{(3)}(s) &= \lambda_4 \cdot s^{(7)} \\
\mu^{(3)}(s) &= \mu^{(3)} \chi(s^{(0)} = 0) \chi(s^{(1)} = 0) \chi(s^{(2)} = 0) \\
p^{(3)}(s) &= 1 \cdot \chi(s^{(0)} = 0) \chi(s^{(1)} = 0) \chi(s^{(2)} = 0) \chi(s^{(10)} = R_4)
\end{aligned} \tag{V.4}$$

$$\begin{aligned}
\lambda^{(j)}(s) &= \alpha_{\text{OFF}}^{(j-4)} (N_j - s^{(j)}), \quad j = 4, 5, 6, 7 \\
\mu^{(j)}(s) &= \alpha_{\text{ON}}^{(j-4)} s^{(j)}, \quad j = 4, 5, 6, 7
\end{aligned} \tag{V.5}$$

$$\begin{aligned}
\lambda^{(8)}(s) &= \mu_R^{(8)} \chi(s^{(0)} \neq 0) \\
p^{(8)} &= 1 \cdot \chi(s^{(0)} \neq 0)
\end{aligned} \tag{V.6}$$

$$\begin{aligned}
\lambda^{(9)}(s) &= \mu_R^{(9)} \chi(s^{(0)} = 0) \chi(s^{(1)} \neq 0 \vee s^{(2)} \neq 0) \\
p^{(9)} &= 1 \cdot \chi(s^{(1)} \neq 0 \vee s^{(2)} \neq 0)
\end{aligned} \tag{V.7}$$

$$\begin{aligned}
\lambda^{(10)}(s) &= \mu_R^{(10)} \chi(s^{(0)} = 0) \chi(s^{(1)} = 0) \chi(s^{(2)} = 0) \chi(s^{(3)} \neq 0) \\
p^{(10)} &= 1 \cdot \chi(s^{(0)} = 0) \chi(s^{(1)} = 0) \chi(s^{(2)} = 0) \chi(s^{(3)} \neq 0)
\end{aligned} \tag{V.8}$$

Grâce à la méthode les automates stochastiques nous avons été capables de construire des modèles markoviens de grand taille d'un système complexe. L'introduction de transitions fonctionnelles facilite considérablement la compréhension et la construction de la matrice de transitions markovienne. Les solutions de ces deux modèles ainsi que les résultats de simulation sont dans [TEA00].

2 Switch tout-optique

Nous avons proposé des modèles markoviens en temps discret pour des trois architectures de commutateur tout-optique présentées dans le paragraphe 2 du chapitre II et nous avons calculé le ratio de paquets perdus (*Packet Loss Ratio*, PLR) comme mesure de performance.

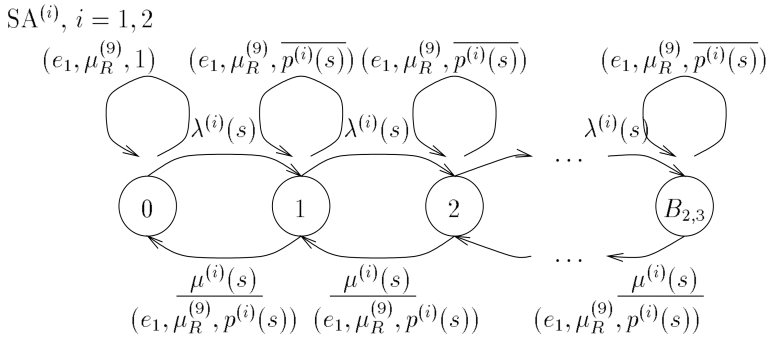

Fig. V.2: *Automate $SA^{(1)}$ représentant le nombre de paquets de vidéo et automate $SA^{(2)}$ représentant le nombre de paquets de voix, Eq. (V.2), (V.3)*

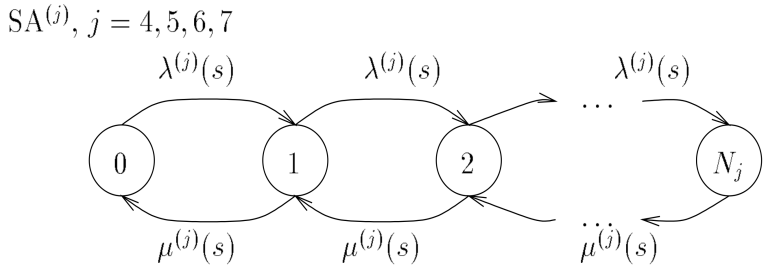

Fig. V.3: *Automates $SA^{(j)}, j = 4, 5, 6, 7$ représentant le nombre de sources ON/OFF de paquets de classe $j - 3$ actives, MMPP, Eq. (V.5)*

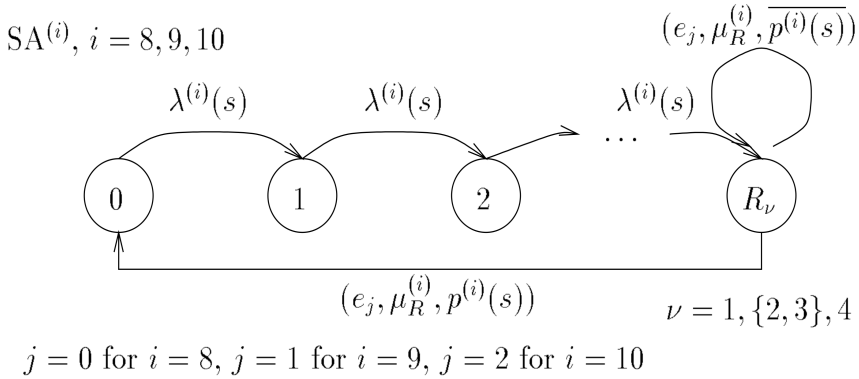

Fig. V.4: *Automates $SA^{(j)}, j = 8, 9, 10$ approximant le temps de traitement constant, Eq. (V.6), (V.7), (V.8)*

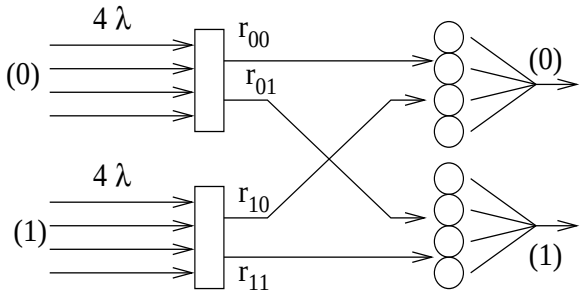

Fig. V.5: *Modèle du commutateur sans CRM*

Pour décrire la chaîne correspondant à l'architecture sans CRM nous nous sommes servis du modèle illustré par Figure V.5 (deux fibres à $N = 4$ longueurs d'onde chacune) dans lequel le temps de traitement de tout paquet est égal à sa longueur. La matrice $R = [r_{ij}]_{i,j=0,1}$, $\sum_j r_{ij} = 1$ définit les probabilités de passage entre les fibres (r_{ij} indique la probabilité qu'un paquet provenant de la fibre (i) est dirigé vers la fibre (j), $i, j = 0, 1$). Dans cette architecture, si un paquet ne peut pas entrer immédiatement dans la fibre choisie, il sera perdu.

Nous avons supposé qu'une longueur d'onde de la fibre (i), $i = 0, 1$ est occupée avec probabilité $p^{(i)}$ qui peut également comprise comme la charge moyenne d'une longueur d'onde. La distribution de l'occupation d'une longueur d'onde peut être donc décrite par une distribution de Bernouilli avec la probabilité de succès $p^{(i)}$ et la probabilité d'échec $q^{(i)} = 1 - p^{(i)}$. La supposition de l'indépendance des longueurs d'onde nous a permis de représenter la distribution de nombre de longueurs d'onde occupées dans une fibre par une distribution binomiale. Après avoir noté $p_n^{(i)}$ la probabilité que n longueurs d'onde sont prises dans la fibre i, nous avons donc :

$$p_n^{(i)} = \binom{N}{n} p^{(i)^n} (1 - p^{(i)})^{N-n}. \tag{V.9}$$

La charge moyenne de la fibre est la même que celle d'une longueur d'onde car la moyenne de la distribution donnée par Eq. (V.9) est de $Np^{(i)}$.

Un état de la chaîne de Markov décrivant le comportement de ce commutateur est composé de quatre éléments : $\vec{s} = \left(s^{(0)}, s^{(1)}, s^{(2)}, s^{(3)}\right)$ où

- $s^{(0)}$ — nombre de longueurs d'onde prises dans la fibre (0),
- $s^{(1)}$ — nombre de longueurs d'onde prises dans la fibre (1),
- $s^{(2)}$ — nombre de places prises dans serveur (0),
- $s^{(3)}$ — nombre de places prises dans serveur (1)

et le nombre de ses états est égal à $(N + 1)^4$.

Pour trouver la matrice stochastique de la chaîne de Markov décrivant ce modèle, nous calculons les probabilités de transitions entre un état courant $\vec{s}$ et le suivant $\vec{s}_*$ marquées comme $p(\vec{s}, \vec{s_*})$. Nous observons que $s_*^{(0)}$, $s_*^{(1)}$ ne dépendent pas du tout de $\vec{s}$ et $s^{(2)}$, $s^{(3)}$ n'ont aucune influence sur $\vec{s}_*$. Par conséquent,

$$p(\vec{s}, \vec{s_*}) = p^{(0)}_{s_*^{(0)}} p^{(1)}_{s_*^{(1)}} \mathrm{pr}(\vec{s}, \vec{s_*}), \tag{V.10}$$

où $p^{(i)}_{s_*^{(i)}}$, $i = 0, 1$, est la probabilité que $s_*^{(i)}$ longueurs d'onde sont occupées dans la fibre (i) et $\mathrm{pr}(\vec{s}, \vec{s_*})$ est la probabilité de routage de paquets arrivants vers une sortie. Nous trouvons la probabilité de

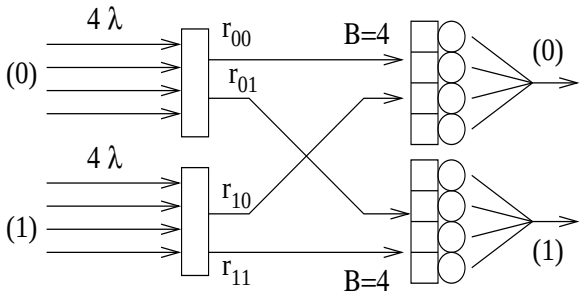

Fig. V.6: *Modèle du commutateur avec lignes à retard*

routage de ci-dessus par une formule combinatoire :

$$\mathrm{pr}(\vec{s},\vec{s_*}) = \sum_{k=0}^{h} \binom{s^{(0)}}{h-k} \binom{s^{(1)}}{k} r_{00}^{h-k} r_{01}^{s^{(0)}-(h-k)} r_{10}^{k} r_{11}^{s^{(1)}-k} \tag{V.11}$$

dans laquelle nous mettons $\binom{n}{k} = 0$ si $k > n$. Pour déterminer l'indice de somme h, nous considérons des paquets qui ont été routés mais qui sont, peut-être, perdus. Les pertes se produisent quand le nombre de paquets entrants dépasse le nombre de ceux qui partent pendant un créneau suivant $s^{(0)} + s^{(1)} > s_*^{(2)} + s_*^{(3)}$. Pour cette raison nous considérons $h = s^{(0)} + s^{(1)} - s_*^{(3)}$. Dans le cas sans pertes $h = s_*^{(2)}$. Dans [TK09] nous généralisons Eq. (V.10) et (V.11) pour les fibres de M longueurs onde.

Le modèle du système précédent est maintenant adapté pour représenter le commutateur équipé d'une mémoire optique de taille $B = 4$ (Figure V.6) qui est utilisée dans le cas où les serveurs sont occupés. Cette adaptation est faite par le changement de structure de l'état de la chaîne de Markov qui doit tenir compte de l'existence de tampons : $\vec{s} = \left(s^{(0)}, s^{(1)}, s^{(2)}, s^{(3)}, s^{(4)}, s^{(5)}\right)$ où

$s^{(0)}$, $s^{(1)}$ — nombre de longueurs d'onde prises respectivement dans la fibre entrante (0) et (1),
$s^{(2)}$, $s^{(3)}$ — nombre de places prises respectivement dans les tampons (0) et (1),
$s^{(4)}$, $s^{(5)}$ — nombre de longueurs d'onde prises respectivement dans la fibre sortante (0) et (1).

Le calcul du nombre d'état de cette chaîne de Markov nécessite l'analyse de quatre cas possibles :

1. $(s^{(0)}, s^{(1)}, 0, 0, s^{(4)}, s^{(5)})$ — les tampons ne sont pas utilisés, toutes les valeurs de $s^{(j)}$, $j = 0, 1, 4, 5$ sont possibles, il y a donc $(N+1)^4$ de tels états,
2. $(s^{(0)}, s^{(1)}, \{1, 2, \ldots, B\}, 0, N, s^{(5)})$ — il y a au moins un paquet dans le tampon (0) ; par conséquent, le serveur (0) voit toutes ses places prises, le tampon (1) est vide ; donc le nombre de ces états est $B(N+1)^3$,
3. $(s^{(0)}, s^{(1)}, 0, \{1, 2, \ldots, B\}, s^{(4)}, N)$ — c'est une situation duale de la précédente : $B(N+1)^3$ d'états,
4. $(s^{(0)}, s^{(1)}, \{1, 2, \ldots, B\}, \{1, 2, \ldots, B\}, N, N)$ — au moins un paquet se trouve dans chacun des tampons, alors les deux serveurs sont entièrement occupés. Comme nous le démontrons dans [TKA03], puisque la somme des nombres de paquets attendant dans deux tampons ne peut pas dépasser la capacité d'un tampon ($s^{(2)} + s^{(3)} \leq B$), le nombre d'états de cette catégorie est égal à $(N+1)^2 \frac{B(B-1)}{2}$.

Le nombre d'état de la chaîne est donc la somme des nombres d'états de ces groupes

$$(N+1)^2 \left[(N+1)^2 + 2B(N+1) + \frac{B(B-1)}{2} \right].$$

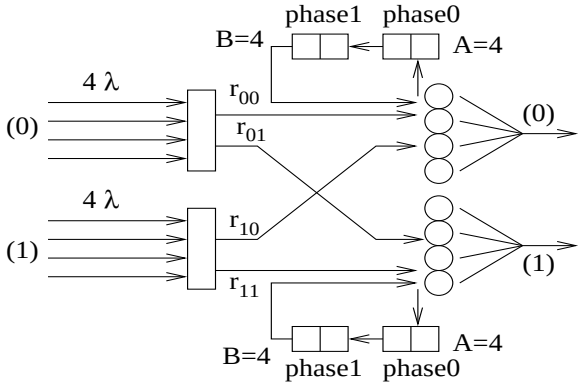

Fig. V.7: *Modèle du commutateur avec* feed-back

Les éléments de la matrice stochastique de la chaîne sont trouvés selon les formules (V.10) et (V.11).

La solution consistant à introduire le mécanisme *feed-back* entre les ports de sortie et les ports d'entrée est réalisée par des lignes à retard connectées en cascade (Figure V.7). Si un paquet ne peut pas entrer dans un serveur, il est temporairement stocké sur un premier niveau de lignes à retard dont la capacité exprimée en nombre de paquets est A. Ensuite, lors d'un créneau suivant, il est déplacé vers le deuxième niveau de lignes à retard dont la capacité est B. Après avoir subit deux fois le délai, le paquet est redirigé vers un serveur de sortie.

La structure d'un état de la chaîne de Markov est dans ce cas la suivante : $\vec{s} = \left(s^{(0)}, s^{(1)}, s^{(2)}, s^{(3)}, s^{(4)}, s^{(5)}, s^{(6)}, s^{(7)}\right)$ où

$s^{(0)}$, $s^{(1)}$ — nombre de paquets arrivant respectivement sur la fibre (0) et (1),

$s^{(2)}$, $s^{(3)}$ — nombre de paquets en dernière phase de *feed-back*, qui vont réessayer de trouver un serveur disponible, respectivement sur les fibres (0) et (1),

$s^{(4)}$, $s^{(5)}$ — nombre de longueurs d'onde prises respectivement dans les fibres sortantes (0) et (1) (occupation des serveurs),

$s^{(6)}$, $s^{(7)}$ — nombre de paquets en dernière phase de *feed-back*, respectivement sur les fibres (0) et (1).

Avant de calculer le nombre d'états de la chaîne, nous observons que le contenu de lignes à retard du premier niveau est déplacé vers le niveau supérieur dans l'état suivant : $s_*^{(2)} = s^{(6)}$ et $s_*^{(3)} = s^{(7)}$. Nous constatons également que le nombre de paquets présent dans un serveur est supérieur ou égal au nombre de paquets au deuxième niveau de lignes à retard de l'état précédent : $s_*^{(4)} \geq s^{(2)}$ et $s_*^{(5)} \geq s^{(3)}$. De plus, les limites sur les nombres de paquets se trouvant au même niveau de lignes à retard restent les mêmes que dans le modèle avec mémoire. Nous analysons trois groupes d'états possibles :

1. Quand les deux lignes à retard de deuxième niveau sont vides ($s^{(2)} = s^{(3)} = 0$), alors nous réutilisons les résultats pour un de groupes du modèle précédent en obtenant

$$n_1 = (N+1)^2 \left[\frac{A(A+1)}{2} + (N+1)(2A+N+1)\right].$$

2. Quand les deux groupes de lignes à retard contiennent au moins un paquet, nous tenons compte du fait que $s^{(2)} + s^{(3)} \leq B$ et par conséquence $n_2 = \frac{B(B-1)}{2} \cdot n_1$.
3. Quand un groupe de lignes à retard est vide et que l'autre contient au moins un paquet, le nombre d'états correspondant est égal à $n_3 = 2Bn_1$.

Le nombre d'états de la chaîne modélisant le système avec *feed-back* est donc

$$n = \left(1 + 2B + \frac{B(B-1)}{2}\right) n_1.$$

Les éléments de la matrice stochastique de la chaîne suivent toujours les formules (V.10) et (V.11).

L'article [TK09] contient les résultats des solutions des chaînes de Markov grâce auxquelles nous avons calculé le taux de perte de paquets selon la formule

$$\mathrm{PLR} = \frac{\sum_{(\vec{s},\vec{s_*})} \Delta(\vec{s},\vec{s_*}) \vec{P}_{(\vec{s},\vec{s_*})} \vec{\pi}_{\vec{s}}}{N(\text{charge moyenne pour la fibre (0)} + \text{charge moyenne pour la fibre (1)})}.$$

Les calculs de $\Delta(\vec{s},\vec{s_*})$ dépendent du modèle. Pour le modèle sans CRM nous avons : $\Delta(\vec{s},\vec{s_*}) = s^{(0)} + s^{(1)} - (s_*^{(2)} + s_*^{(3)})$, pour le modèle avec mémoire optique : $\Delta(\vec{s},\vec{s_*}) = s^{(0)} + s^{(1)} + s^{(2)} + s^{(3)} - (s_*^{(2)} + s_*^{(3)} + s_*^{(4)} + s_*^{(5)})$ et pour le modèle avec lignes *feed-back* : $\Delta(\vec{s},\vec{s_*}) = s^{(0)} + s^{(1)} + s^{(2)} + s^{(3)} - (s_*^{(4)} + s_*^{(5)} + s_*^{(6)} + s_*^{(7)})$. Les détails d'adaptation de cette formule pour le cas général, c'est à dire$M$ fibres composées de N longueurs d'onde, sont présentés dans [TK09].

Les résultats analytiques ont également été confrontés avec les estimations obtenues grâce à la simulation. De plus, les modèles discutés ont été enrichis par l'introduction de chaînes de Markov modulantes qui permettent de varier la charge sur les fibres d'entrée. L'état d'une chaîne modulante pour une fibre décrit le nombre de sources ON/OFF identiques et indépendantes qui sont actives et qui y émettant des paquets. Le coût de cette modification se chiffre par l'ajout des éléments pour pouvoir déterminer la probabilité de la charge à l'entrée du commutateur, par conséquent, par l'augmentation de nombre d'états d'une chaîne de Markov globale. Si une chaîne modulante la charge pour i-ième fibre, $i = 1, 2, \ldots, M$ est de taille $m^{(i)}$ alors le facteur d'augmentation est de $\prod_{j=1}^{M} m^{(i)}$.

3 Accès au réseau optique

Dans ce paragraphe nous présentons les modèles de la politique d'accès CSMA/CA au réseau optique en double anneau DBORN sur le bus d'écriture (*upstream bus*) dont l'architecture et le contexte technologique a été décrit en détail dans le paragraphe 3 du chapitre II. Nous avons traité individuellement les nœuds $1, 2, \ldots, N$ en tenant compte de l'influence des nœuds précédant un nœud courant sur l'occupation de la fibre.

3.1 Paquets optiques de taille variable

Le schéma du modèle est représenté sur Figure V.8. Pour modéliser le trafic accédant au réseau DBORN asynchrone nous avons supposé que les nœuds de l'anneau injectent sur la fibre des paquets IP. Suivant les analyses fournies par [CAI02] nous avons distingué trois classes de paquets en fonction de leur longueur. Les paquets les plus courts sont composés de 50 octets (par exemple TCP/IP ACK) et ils constituent 10% du trafic global. Cette quantité de données, 50 octets, sera vue comme la plus petite portion de données et notée comme u. La deuxième classe contient des paquets de taille moyenne de 500 octets, ($10u$) et représente 40% du volume de trafic. La troisième classe contient les paquets les plus grands dont la longueur est égale à la longueur du paquet Ethernet le plus long (MTU Ethernet), c'est à dire de 1500 octets ($30u$) et constitue 50% de volume de trafic restant. Pour modéliser ces trois flux de paquets différent nous avons utilisé des sources poissoniennes dont les paramètres sont respectivement égaux à γ_1, γ_2 et γ_3.

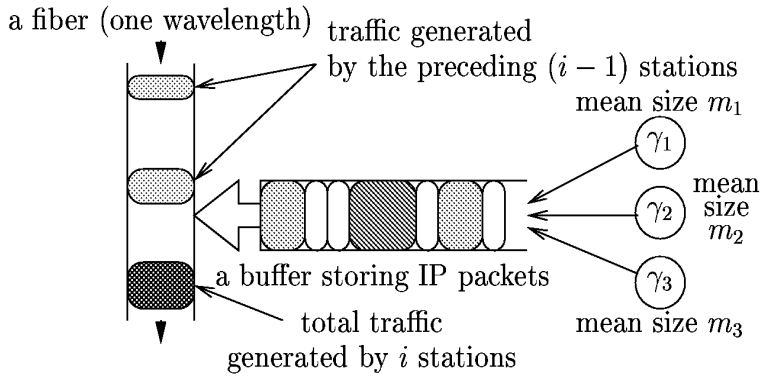

Fig. V.8: *Schéma d'accès au médium physique sur le bus d'écriture dans un nœud i, paquet optique de taille variable*

La priorité due à la position du nœud implique que le nœud le plus proche du concentrateur dispose toujours de la bande passante. Les nœuds de numéro i, $i = 2, \ldots, N$ observent le médium physique en permanence. Quand le canal optique est occupé par des paquets envoyés par les prédécesseurs du nœud courant, les paquets à envoyer attendent dans un tampon électronique. Quand le canal optique devient libre, un nœud commence à mesurer la longueur d'un créneau disponible. Nous notons la durée de transmission de l'unité u de données t_u. Suite au classement de paquets par rapport à leur taille nous pouvons distinguer trois longueurs de créneau à mesurer, t_u, $10t_u$ et $30t_u$, nécessaires pour insérer sans collision respectivement un paquet de première, deuxième et troisième classe. Dans le cas où un créneau disponible pour un nœud i est plus long que la durée de la fenêtre d'observation égale à la durée de transmission des paquets les plus longs (c-à-d $30t_u$) ce nœud remet son compteur de temps à zéro.

La chaîne de Markov modélisant un nœud du réseau est en temps continu et son état se compose de quatre éléments : $(s^{(0)}, s^{(1)}, s^{(2)}, s^{(3)})$. La coordonnée $s^{(0)}$ indique l'état du canal optique et la longueur d'un créneau disponible pour un nœud i. Les coordonnées suivantes indiquent les nombres de paquets de classes 1, 2 et 3 attendant dans le tampon électronique FIFO du nœud i. Afin de garder le nombre d'états de la chaîne au-dessous de la limite raisonnable pour les calculs, l'ordre selon lequel les paquets sont stockés dans le tampon ne fait pas partie d'état de la chaîne. La probabilité p_k qu'un paquet de classe $k = 1, 2, 3$ se trouve en tête du tampon et qu'il sortira le premier est déterminée par la formule

$$p_k = \frac{s^{(k)}}{\sum_{j=1}^{3} s^{(j)}}. \tag{V.12}$$

Cette chaîne de Markov est construite grâce au formalisme des Réseaux d'Automates Stochastiques : un automate $\mathrm{SA}^{(i)}$ gère le comportement de l'élément $s^{(i)}$ de l'état markovien. A cause de la situation privilégiée du nœud le plus proche du concentrateur, nous proposons une version de $\mathrm{SA}^{(0)}$ particulière pour lui.

Nous commençons la description du réseau d'automates stochastiques décrivant le point d'accès à la fibre optique par l'automate de la fibre $\mathrm{SA}^{(0)}$ pour les nœuds i différents du plus proche du concentrateur (Figure V.9), ceux alors qui peuvent voir des créneaux occupés. L'état 0 correspond à l'état occupé du canal optique, un état j, $j = 1, 2, \ldots, 31$ signifie qu'un nœud a observé un créneau suffisant pour y émettre $(j-1)u$. Le canal optique devient occupé avec le taux de transition

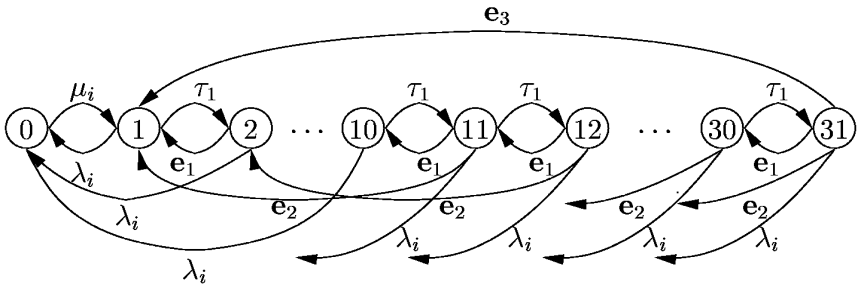

FIG. V.9: *$SA^{(0)}$: état du canal optique vu par les nœud i, $i = 2, 3, \ldots, N$*

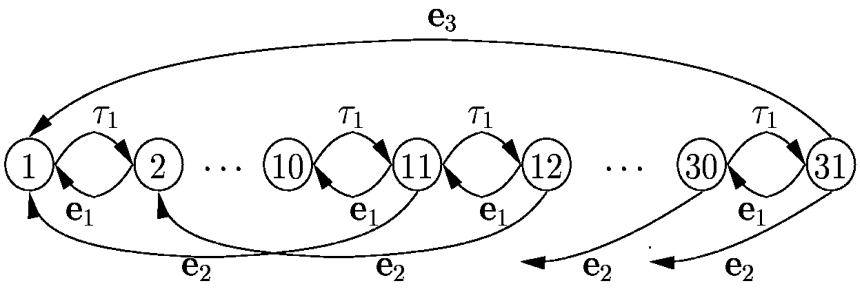

FIG. V.10: *$SA^{(0)}$: état du canal optique vu par le nœud le plus proche du concentrateur*

locale exponentiel λ_i et il se libère par la transition exponentielle locale de taux μ_i. Ces transitions gérant l'état de la fibre ne sont pas présentes dans l'automate du premier nœud (le plus proche du concentrateur) car pour lui le canal est toujours disponible (Figure V.10).

L'unité de temps pour effectuer des mesures est représentée par la transition locale exponentielle dont le taux τ_1 correspond au temps de transmission d'un plus court paquet ($\tau_1 = \frac{1}{t_u}$). L'automate $\mathrm{SA}^{(0)}$ doit collaborer avec les automates chargés de garder les nombres de paquets électroniques attendant la transmission dans un tampon au moment de la transmission d'un paquet. Cette collaboration est réalisée par les événements synchronisants e_k, $k = 1, 2, 3$ indiquant un départ de paquet de classe k, qui ont lieu en même temps dans les automates $\mathrm{SA}^{(0)}$ et $\mathrm{SA}^{(k)}$. La description complète des e_k, illustré par Figure V.11, est donnée par un triplet $(e_k, \theta_k, p_k\mathcal{C}_k)$, où θ_k est un taux de transition correspondant au temps nécessaire au le protocole MAC nécessite pour détecter l'état du canal (quelques nanosecondes), p_k est la probabilité qu'un paquet de classe k soit en tête du tampon (Eq. (V.12)) et $\mathcal{C}_k$ est la condition de déclenchement d'événement. Cette condition peut avoir une valeur différente de 1 uniquement pour e_3 où elle sert à distinguer deux situations dans lesquelles cet événement peut se produire : soit un paquet de classe 3 est transmis, soit une fenêtre d'observation se termine sans paquets à envoyer (dans ce cas nous fixons $p_3 = 1$, voir la boucle pointillée sur Figure V.11).

Afin de connaître le nombre d'états de la chaîne de Markov nous calculons le nombre de configurations de paquets possibles dans le tampon électronique. Supposons que la taille de ce tampon soit de Lu et notons les longueurs des paquets par $\alpha_1 u$, $\alpha_2 u$, $\alpha_3 u$ ($\alpha_1 = 1$, $\alpha_2 = 10$, $\alpha_3 = 30$). Le nombre de paquets de chaque classe qui peuvent être stockés dans le tampon vide est égal à $B_i = L/\alpha_i$, $i = 1,2,3$. Les paquets stockés dans le tampon ne dépassant pas sa capacité, nous avons

$$\alpha_1 s^{(1)} + \alpha_2 s^{(2)} + \alpha_3 s^{(3)} \leq L. \tag{V.13}$$

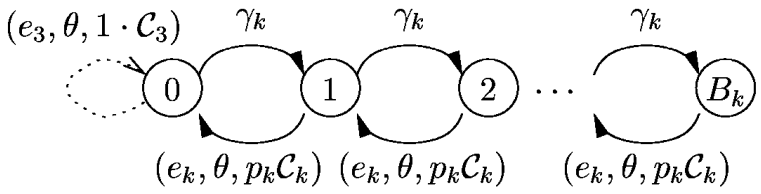

Fig. V.11: *$SA^{(k)}$, $k = 1, 2, 3$: le nombre de paquets de la classe k dans un nœud i*

Si seuls des paquets de première classe sont présents dans le tampon ($s^{(0)} \neq 0$), le nombre de situations possibles (le tampon vide inclus) est donné par

$$\mathcal{N}_1(L) \;=\; B_1 + 1 \;=\; L + 1.$$

Pour des paquets de classe 1 et 2 présents dans le tampon, le nombre de configurations peut être trouvé grâce à la formule :

$$\begin{array}{rcl} \mathcal{N}_2(L) & = & \sum_{i=0}^{B_2} \mathcal{N}_1(L - i\alpha_2) \\ & = & (L+1)(B_2+1) - \frac{1}{2}\alpha_2 B_2(B_2+1) \\ & = & (\frac{L}{\alpha_2}+1)(\frac{L}{2}+1). \end{array}$$

Poursuivant le même raisonnement pour les paquets des trois classes, nous avons

$$\mathcal{N}_3(L) = \sum_{i=0}^{B_3} \mathcal{N}_2(L - i\alpha_3) \tag{V.14}$$

et

$$\mathcal{N}_m(L) = \sum_{i=0}^{B_m} \mathcal{N}_{m-1}(L - i\alpha_m)$$

dans le cas de m classes de paquets. Pour notre modèle le nombre d'états de la chaîne globale est $\mathcal{N}(L) = 32\mathcal{N}_3(L)$ car le nombre d'états de l'automate $\mathrm{SA}^{(0)}$ est égal à 32.

Le modèle markovien décrit ci-dessus nous a servi à calculer le délai moyen d'accès au médium physique et la probabilité de perte de paquets pour un canal de 2.5 Gbps avec $N = 8$ nœuds d'accès identiques notamment au niveau de leur émission ($\gamma^i = \sum_{k=1}^{3} \gamma_k$, $i = 1, \ldots, N$, $\gamma^i = \gamma^j = \gamma$ $(i \neq j)$. Nous avons résolu ce modèle pour deux tailles de tampon de nœud, $L = 60$ et $L = 300$. L'influence de la position d'un nœud sur la disponibilité du canal a été représentée par la manière de calculer les taux $\lambda_j = (j\text{-}1)\gamma$, $\mu_j = [N\text{-}(j\text{-}1)]\gamma$, $j = 2, 3, \ldots, N$. Les résultats de ce modèle ont été confrontés à ceux obtenus par simulation faite avec NS2 [NS07] et il peuvent être trouvés dans [TP07]. Ils montrent que ces deux méthodes sont complémentaires car le modèle markovien capte mieux les mesures de pertes dans les nœuds proches du concentrateur, mais il ne tiens pas compte du phénomène HoL visible dans la simulation.

3.2 Paquets optiques de taille fixe

Le schéma du modèle est représenté par Figure V.12. La nature synchrone du réseau impose une chaîne de Markov en temps discret.

Comme dans le cas précédent nous nous somme inspirés de [CAI02] afin de déterminer les flots de paquets électroniques arrivant dans un nœud. Nous avons pris les taux de génération $\gamma^{(k)}$, $k =$

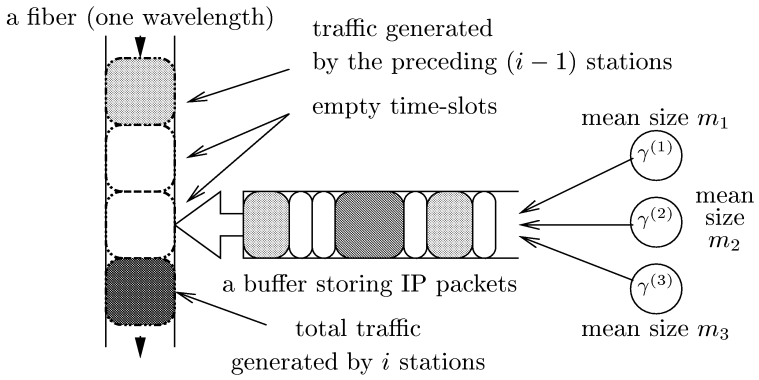

Fig. V.12: *Schéma d'accès au médium physique sur le bus d'écriture dans un nœud i, paquet optique de taille fixe*

$1, 2, 3$ pour calculer la probabilité d'arrivée de l paquets d'une classe donnée pendant la durée d'un paquet optique dont la longueur exprimée en unités de temps est égale à τ : $p_l^{(k)}(\tau) = \frac{\alpha^{(k)l} e^{-\alpha^{(k)}}}{l!}$ où $\alpha^{(k)} = \gamma^{(k)}\tau$.

L'état de la chaîne de Markov pour un nœud i est de la forme $s = (s^{(0)}, s^{(1)}, s^{(2)}, s^{(3)})$. Le composant $s^{(0)}$ indique l'état du canal optique ($s^{(0)} = 0$ quand un paquet optique couramment observe est occupé et $s^{(0)} = 1$ quand il est libre). Les composants $s^{(k)}$, $k = 1, 2, 3$ représentent les nombres de paquets de classe k attendant dans le tampon de ce nœud. Comme dans le cas asynchrone, nous ne gardons pas l'ordre des paquets électroniques afin de réduire le nombre d'états de la chaîne. Les états de la chaîne peuvent être ordonnés lexigraphiquement et numérotés. La matrice $P = [p_{ij}]$ où p_{ij} est la probabilité de transition de l'état indexé par i à l'état indexé par j est la matrice stochastique définissant la chaîne de Markov. La difficulté consiste à trouver les valeurs p_{ij}.

Le nombre d'états de cette chaîne de Markov est calculé d'une manière semblable à celle du cas asynchrone. Nous commençons par poser la même inégalité (V.13) et nous donnons la formule déterminant le nombre de compositions dans le tampon comme Eq. (V.14). Le nombre d'états pour un nœud i, avec $i \neq 1$ est donc égal à $\mathcal{N}(B) = 2\mathcal{N}_3(B)$. Pour le premier nœud le nombre d'états est simplement égal à $\mathcal{N}_3(B)$ car ce nœud a toujours accès au canal et le composant $s^{(0)}$ peut être retiré.

Pour construire la matrice P les événements sont observés à la fin de chaque créneau de temps. L'observateur extérieur d'un nœud i voit donc le canal dont occupation est déterminée par les émissions de $i-1$ nœuds précédents au moment où tous les paquets électroniques présents au début de chaque créneau et ceux qui sont arrivés durant ce créneau et dont la capacité totale ne dépasse pas la longueur d'un paquet optique sont déjà partis (à condition que le canal ait été disponible pour ce nœud).

Nous supposons que tout les nœuds sont identiques et que la probabilité qu'un nœud utilise un paquet optique courant est égale à p. Un nœud i voit donc le canal occupé avec la probabilité $(i-1)p$. La valeur p correspond aussi à la fraction de la bande passante du canal prise en moyenne par chaque nœud. Après avoir noté $\vec{s}$ un état courant et $\vec{s_*}$ un état suivant et avoir observé que l'état

du canal ne dépend pas d'un nœud courant nous écrivons :

$$\begin{array}{l} P\left(\vec{s},(0,s_*^{(1)},s_*^{(2)},s_*^{(3)})\right) = \\ (i-1)pP^*\left((s^{(1)},s^{(2)},s^{(3)}),(s_*^{(1)},s_*^{(2)},s_*^{(3)})\right) \\ P\left(\vec{s},(1,s_*^{(1)},s_*^{(2)},s_*^{(3)})\right) = \\ (1-(i-1)p)P^*\left((s^{(1)},s^{(2)},s^{(3)}),(s_*^{(1)},s_*^{(2)},s_*^{(3)})\right) \end{array} \tag{V.15}$$

où $P^*\left((s^{(1)},s^{(2)},s^{(3)}),(s_*^{(1)},s_*^{(2)},s_*^{(3)})\right)$ est la probabilité de changement d'état du tampon $(s^{(1)}, s^{(2)}, s^{(3)}) \longrightarrow (s_*^{(1)}, s_*^{(2)}, s_*^{(3)})$ et sera notée comme $P^*(\tilde{\mathbf{s}}, \tilde{\mathbf{s_*}})$ dans la suite.

Les calculs de $P^*(\tilde{\mathbf{s}}, \tilde{\mathbf{s_*}})$ doivent tenir compte des arrivées de paquets électroniques pendant une période égale à la longueur d'un paquets optique et du départ éventuel d'un paquet optique complet, à condition que ce départ puisse se produire (à condition que $s^{(0)} = 1$). Nous limitons les nombres de paquets électroniques de chaque classe k arrivant dans un nœud pendant la durée d'un paquet optique par $n^{(k)}$, $k = 1, 2, 3$. Nous notons $p_{l^{(k)}}$ que $l^{(k)}$ paquet de classe k arrivent durant cette période. Vu la nature discrète du système, des paquets électroniques arrivent par lot, $l = (l^{(1)}, l^{(2)}, l^{(3)})$ et par conséquence la probabilité d'arrivée un lot l est égale à $\pi_l = p_{l^{(1)}} p_{l^{(2)}} p_{l^{(3)}}$ car les arrivées de paquets électroniques sont indépendantes. N'ayant pas d'informations sur l'ordre de paquets arrivants, nous calculons le nombre de compositions possibles d'un lot comme $n_{\mathbf{l}} = \binom{l^{(1)}+l^{(2)}+l^{(3)}}{l^{(1)}} \binom{l^{(2)}+l^{(3)}}{l^{(2)}}$. Toutes les compositions d'un lot sont équiprobables. L'ordre de paquets dans un lot est important pour déterminer les situations de dépassement de la capacité du tampon et la composition d'un paquet optique à transmettre.

Nous supposons qu'un lot l arrive au moment où un état courant du nœud i est $\vec{s}$ et où le canal optique n'est pas disponible ($s^{(0)} = 0$). Dans le cas où le tampon peut absorber tous les paquets arrivant dans ce lot, l'état suivant est $\vec{s_*} = (0/1, s^{(1)} + l^{(1)}, s^{(2)} + l^{(2)}, s^{(3)} + l^{(3)})$ et

$$P^*\left(\tilde{\mathbf{s}}, (s^{(1)} + l^{(1)}, s^{(2)} + l^{(2)}, s^{(3)} + l^{(3)})\right) = \pi_{\mathbf{l}}. \tag{V.16}$$

L'autre cas étudié a lieu quand la place libre dans le tampon, notée comme F unités u, n'est pas suffisante pour stocker un lot l. Nous représentons une configuration d'un lot possible comme $c_l = (c_1^{(k_1)}, c_2^{(k_2)}, \ldots, c_l^{(k_l)})$ où $c_j^{(k_j)}$ est pour paquet électronique en position j. Pour une composition d'un lot donnée c_l donné nous procédons au fur et à mesure acceptant dans le tampon $c_j^{(k_j)}$ paquets quand $F - c_j^{(k_j)} \geq 0$ et recommençant après avoir diminué la valeur de F. Nous notons comme $m^{(k)}, k = 1, 2, 3$ un nombre de paquets de classe k acceptés pour une configuration c_l et $m = \sum_{k=1}^{3} m^{(k)}$. L'ensemble $\mathcal{L}_m$ est constitué de tous les lots l ayant des configurations permettant l'acceptation de m paquets électroniques dans le tampon dans l'état s d'un nœud i. L'ensemble $L(\mathrm{l}, m)$ regroupe toutes les configuration d'un lot l pour lesquelles m paquets de ce lot est accepté dans le tampon dans l'état s d'un nœud. La probabilité de transition vers l'état suivant $\vec{s_*} = (0/1, s^{(1)} + m^{(1)}, s^{(2)} + m^{(2)}, s^{(3)} + m^{(3)})$ est dons exprimée par

$$\begin{array}{l} P^*\left(\tilde{\mathbf{s}}, (s^{(1)} + m^{(1)}, s^{(2)} + m^{(2)}, s^{(3)} + m^{(3)})\right) = \\ \sum_{\mathbf{l} \in \mathcal{L}_m} \frac{1}{n_{\mathbf{l}}} \pi_{\mathbf{l}} |L(\mathrm{l}, m)|. \end{array} \tag{V.17}$$

Nous supposons maintenant, qu'au moment de l'arrivée d'un lot, le canal optique est disponible ($s^{(0)} = 1$). Nous permettons la transmission d'un paquet optique uniquement quand il peut être totalement remplie. Ainsi dans les situations dans lesquelles des paquets électroniques attendant

dans le tampon et ceux qui arrivent ne remplissent pas un paquet, la probabilité de transition vers un état suivant est exprime par Eq. (V.16). Considérons maintenant les cas où des paquets électroniques présents dans le tampon peuvent constituer un paquet optique complet $I = (i^{(1)}, i^{(2)}, i^{(3)})$. Un lot arrivant a donc à sa disposition l'espace libre du tampon augmenté par la taille de paquet optique que formeront des paquets attendant dans le tampon en le quittant. Le calcul de la probabilité de transition vers un état suivant est similaire à Eq. (V.17) et est écrit comme

$$\begin{array}{l} P^*\left(\overline{s}, (s^{(k)} + m^{(k)} - i^{(k)})_{k=1,2,3}\right) = \\ \sum_{I \in \mathcal{I}} \mathcal{P}_I \sum_{l \in \mathcal{L}_m} \frac{|L(l,m)|}{n_1} \pi_l. \end{array} \tag{V.18}$$

Le cas le plus compliqué se produit quand des paquets attendant dans le tampon ne sont pas suffisamment nombreux pour remplir totalement un paquet optique mais des paquets électroniques arrivant dans un lot peuvent le compléter. Pour déterminer la complétion d'un paquet nous appliquant d'abord la même procédure de considérer de configurations possibles de lot qu'auparavant. Si les éléments d'un lot, après avoir complété un paquet optique, ont assez d'espace libre dans le tampon, nous procédons selon la formule Eq. (V.18). Dans le cas contraire, nous devons calculer la probabilité d'avoir une composition favorable pour former un paquet optique complet. Notons les nombres de paquets électroniques de chaque classe k présents dans le tampon et arrivés pendant la longueur de paquet optique comme $V = (v^{(1)}, v^{(2)}, v^{(3)})$ et $v = \sum_{k=1}^{3} v^{(k)}$. Soit $I = (i^{(1)}, i^{(2)}, i^{(3)})$ une configuration d'un paquet optique complet. Le nombre de combinaisons possibles dans I est donne par $(i^{(1)}, i^{(2)}, i^{(3)})$. Nous observons que le nombre de configurations possibles dans I est $i = \binom{i^{(1)} + i^{(2)} + i^{(3)}}{i^{(1)}} \binom{i^{(2)} + i^{(3)}}{i^{(2)}}$. La probabilité que I puisse être formé de V est calculée comme

$$\mathcal{P}_I = i \cdot \frac{\prod_{k=1}^{3} v^{(k)}(v^{(k)} - 1) \cdots (v^{(k)} - i^{(k)} + 1)}{v(v-1) \cdots (v - i^{(1)}) \cdots (v - (i^{(1)} + i^{(2)} + i^{(3)}) + 1)}.$$

Remarquons qu'avec la probabilité $1 - \mathcal{P}_I$ le départ d'un paquet optique ne se produit pas car l'ordre dans le tampon est défavorable et uniquement l'espace initialement libre est rempli.

Les résultats de ce modèle ainsi que ceux obtenus par simulation ayant comme données les valeurs identiques que celles du cas asynchrone et la longueur de paquet optique est égale à $30u = 4.8\ \mu$ se trouvent dans [TP04]. La supposition faite pour le modèle analytique sur le taux de remplissage de 100% d'un paquet optique augmente les valeurs moyennes des longueurs de tampon dans des nœuds par rapport à celle estimée par simulation qui permet l'envoi d'un paquet incomplet. La discipline de remplissage d'un paquet optique devra donc d'un côté ne pas retenir trop longtemps des paquets dans un nœud et de l'autre côté les faire attendre pour ne pas laisser trop de place inutilisable dans des paquets optiques.

Chapitre VI

Conclusions et perspectives

Tout au long de ce document, nous avons cherché à mettre en évidence notre démarche de recherche au travers d'un certain nombre d'exemples dans le domaine de réseaux de télécommunications et dans le contexte de la garantie de la qualité de service. Notre but était toujours de trouver des modèles de ces réseaux les plus adaptés pour pouvoir estimer différents aspects de leur performance. Ces estimations nous ont permis de détecter des goulots d'étranglement (soit au niveau architectural soit au niveau protocolaire) des systèmes étudiés. Nous avons également proposé des solutions algorithmiques pouvant être implémentées dans de nouveaux protocoles pour améliorer l'efficacité de ces réseaux. Les propriétés technologiques de ces réseaux (technologie tout-optique, technologie radio, paquets optiques de taille fixe et variable, interconnexion de domaines, transmissions multicast) nous ont imposé des approches différentes à la modélisation (analytiques, simulation, temps discret ou continu, générateurs de graphes aléatoires etc.).

Les expériences acquises durant des années de travail sur des thématiques diverses et au sein d'équipes différentes nous permettent de poursuivre de la recherche en explorant de nouvelles pistes que nous semblent être très intéressants. La liste de sujets ci-dessous représente les thématiques de projets potentiels dont la réalisation impliquerait la collaboration d'un groupe de chercheurs, de doctorants et de chercheurs post-doc. Le financement du personnel non-permanent pourrait être assuré par le biais de subventions obtenues dans le cadre de projets scientifiques nationaux et européens.

Réseaux multi-domaines : Nous nous sommes aperçus durant notre travail sur la garantie de la qualité de service dans des réseaux inter-domaines qu'il n'existe pas d'étude fiable permettant d'estimer de temps de transit de trafic par des domaines. La difficulté de la faire est d'un côté la confidentialité des opérateurs qui n'informent pas comment leurs réseaux acheminent le trafic de transit et la difficulté de prédiction de routes établies.

Nous voudrions effectuer une analyse de mesures prises dans Internet afin d'essayer de classer les routes en fonction de leur forme (couche de départ, couche d'arrivée) et trouver une distribution qui pourra être utilisée pour modéliser d'une manière plus fine qu'actuellement le délai introduit par de passages par domaines dans notre générateur aSHIIP. Cette analyse devra aussi nous permettre de détecter des distorsions éventuelles de la hiérarchie provoquées par des changements architecturaux introduits par certains fournisseurs de contenu. Nous voulons également étudier l'importance de liens inter-domaines S2S (*sibling-to-sibling*) qui peuvent exister dans Internet entre des domaines appartenant au même opérateur ou des opérateurs alliés. Nous pensons que ce type de relations entre des domaines peuvent biaiser la structure hiérarchique observée dans Internet et également influencer le temps de transit de trafic.

Réseaux tout-optiques : Les réseaux tout-optiques sont très avantageux dans le contexte du développement écologique car la transmission optique consomme peu d'énergie. Néanmoins, le routage dans ce type de réseaux devint plus contraignant puisque des routeurs tout-optiques ne peuvent pas

créer de copies des paquets qui les traversent. Cette propriété ajoute donc une difficulté supplémentaire à la construction des arbres de transitions multicast. Jusqu'à présent nous avons étudié des placements des arbres de multicast dans des réseaux avec un nombre limité de routeurs pouvant dupliquer les paquets dans le contexte de la bande passante disponible pour une demande de multicast donnée. Il nous semble intéressent d'aborder ce problème dans un cas plus pessimiste dans lequel il faut trouver un arbre dont la bande passante utilisée est minimale parmi les arbres qui assurent la quantité de le ressource réclamé par une demande multicast.

Réseaux sans fil Wi-Fi : Le déploiement d'accès Wi-Fi dans des bâtiments publics et particuliers offre une couverture réseau qui consomme peu d'énergie et qui est potentiellement utilisable à un coût très modéré par des utilisateurs de téléphones mobiles. Nous pensons proposer des algorithmes de décision et de contrôle des *hand-overs* horizontaux (WLAN-WLAN) et verticaux (WLAN-UMTS), basés sur l'anticipation de la mobilité des utilisateurs et sur la prédiction dynamique de la qualité de service que peuvent offrir les différents réseaux d'accès concernés, évaluation et anticipation obtenues par l'analyse du passé.

Cette solution permet de disposer d'une couverture radio partiellement continue de réseaux de collecte très avantageux grâce à la très forte densité du DSL (*Digital Subscriber Line*) et des bornes WiFi en France, et plus généralement en Europe. Ces réseaux pourront être déployés facilement et rapidement par n'importe quel acteur (pas de licence, pas de site à louer). Ils pourront être proposés à des prix très attractifs. Plusieurs motivations fortes pour l'adoption de ce type de réseau existent, du fait de la croissance continue : développement du WiFi, augmentation de nombre de connections mobiles possibles grâce à la décomposition de la même zone UMTS en "plus de" zones WiFi de capacité inférieure à celle de la zone UMTS, pression immobilière sur les sites des antennes et de la pression de l'opinion contre la "pollution radio" (la très faible puissance des points d'accès WiFi peut être mieux perçue que les émissions dans le cellulaire ou le WiMAX).

Réseaux sans fil PMR : Les réseaux PMR (*Private Mobile Radio*) utilisés couramment par des forces de l'ordre (gendarmerie, pompiers, police) sont situés dans le spectre de la bande étroite (*narrow band*). La migration vers des transmissions dans la large bande (*broad band*) est imposée par les utilisateurs qui souhaitent avoir un éventail de services plus riche. Il est prévu que ces réseaux seront contrôlés par la spécification LTE (*Long Term Evolution*). Le passage vers la large bande se heurte au problème de la pénurie de ressources car le spectre disponible est limité. L'application de LTE impose le schéma SC-FDMA (*Single-Carrier FDMA*) sur le lien montant (*uplink*) et OFDMA (*Orthogonal Frequency-Division Multiple Access*) sur le lien descendant (*downlink*). La spécificité de SC-FDMA exige l'allocation contiguë des fréquences et rend les problèmes associés NP-complets. Nous envisageons de travailler sur des nouvelles méthodes de prédiction pour assurer des *hand-overs* efficaces et des automates stochastiques d'apprentissage pour découvrir des valeurs de paramètres de configurations optimales.

Cloud Computing : Suite au développement du réseau Internet et à sa grande accessibilité, les utilisateurs y cherchent des services variés, notamment le stockage. Les fournisseurs de ces services sont indépendants des opérateurs et ils possèdent leurs propres ressources (par exemple les centres de stockage) connectées au réseau.

Nous voyons un intérêt important à développer des méthodes permettant d'ajuster l'infrastructure existante d'un fournisseur de services selon les besoins de ses clients. Cet ajustement a pour but de satisfaire les contrats des clients et de diminuer autant que possible les coûts. Les serveurs fixes sont moins couteux (par heure) que les serveurs "de location" (dont l'heure de location est chère) mais ils exigent un investissement initial, une installation physique et une maintenance. On peut imaginer que le site fixe, quand il commence à être saturé, délègue une partie de sa charge à un site loué ponctuellement. A ce niveau il faut décider quelle partie de service doit être déployée vers un site "de secours". Dans ce contexte il semble judicieux de doter le dispatcher d'un algorithme

"intelligent" d'allocations de ressources pour les services afin de minimiser le nombre de "locations de secours".

Un autre facteur important pour *Cloud Computing* concerne des problèmes environnementaux : la réduction de la consommation énergétique, en conséquence la réduction du coût d'exploitation OPEX (*OPerationnal Expenditure*). Du point de vue des fournisseurs de services (opérateurs de *datacenters*, opérateurs de réseaux, etc) la satisfaction de demandes de leurs clients devra donc être assurée, en particulier la minimisation de l'énergie nécessaire pour alimenter des centres de stockage/calculs d'un côté et la minimisation de l'utilisation de ressources de réseau de l'autre côté. Ces deux axes de minimisation sont à priori indépendants. La minimisation d'énergie consommée par des *datacenters* favorise l'équilibrage de charge biaisé (*skewed load balancing*) pour ne pas "allumer" des unités totalement éteintes, dans le cas de nécessité les allumer en avance car leur mise en marche prend du temps et garder le nombre de centres de services en attente (*stand-by*) raisonnable puisque ces centres dans cet état consomment plus que la moitié d'énergie nominale. La minimisation de la bande passante de réseau souhaitée par un fournisseur de connectivité tente plutôt de réaliser les objectifs d'ingénierie de trafic. Le but des travaux est donc de proposer des méthodes permettant de trouver un point optimal vis-à-vis de ces deux minimisations.

Références bibliographiques

[AB00] R. Albert and A.-L. Barabàsi. Topology of evolving networks : Local events and universality. *Physical Review Letters*, 85 :5234, 2000.

[ACL00] W. Aiello, F. Chung, and L. Lu. A random graph model for massive graphs. In *Proceedings of STOC'00, 32nd annual ACM symposium on theory of computing*, pages 171–180, Portland, Oregon, USA, 2000. ACM Press.

[AGPG+07] N. Aschenbruck, E. Gerhards-Padilla, M. Gerharz, M. Frank, and P. Martini. Modelling mobility in disaster area scenarios. In *Proceedings 10th ACM IEEE International Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems, MSWIM*, 2007.

[Arn51] W. E. Arnoldi. The principle of minimized iterations in the solutions of the matrix eigenvalue problems. *Quart. Appl. Math.*, 19 :17–29, 1951.

[Ati92] K. Atif. *Modellisation du parallelisme et de la synchronisation.* Thèse de l'Institut National Polytechnique de Grenoble, Le Laboratoire de Génie Informatique, September 1992.

[Aud07] O. Audouin. CARRIOCAS description and how it will require changes in the network to support Grids. In *20th Open Grid Forum*, Manchester, UK, 2007.

[BA99] A.-L. Barabàsi and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439) :509, 1999.

[BBC+98] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. RFC 2475 — an architecture for differentiated services, December 1998.

[BBC+05] E.G. Boman, D. Bozdağ, U. Catalyurek, A.H. Gebremedhin, and F. Manne. A scalable parallel graph coloring algorithm for distributed memory computers. *Proceedings of Euro–Par 2005 Parallel Processing*, pages 241–251, 2005.

[BBDP05] N. Bouabdallah, A.-L. Beylot, E. Dotaro, and G. Pujolle. Resolving the fairness issue in bus-based optical access networks. *IEEE JSAC*, 23(8) :1444–1457, August 2005.

[BBF+03] A. Benoit, L. Brenner, P. Fernandes, B. Plateau, and W. J. Stewart. The PEPS Software Tool. In *Computer Performance Evaluation / TOOLS 2003*, volume 2794 of *LNCS*, pages 98–115, Urbana, IL, USA, 2003. Springer-Verlag Heidelberg.

[BBM09] Y. Benallouche, D. Barth, and O. Marcé. Minimizing routing delay variation in case of mobility. In *Proceedings of WiMob'09*, pages 370–375, 2009.

[BCG+05] D. Bozdağ, U. Catalyurek, A.H. Gebremedhin, F. Manne, E.G. Boman, and F. Özgüner. A parallel distance-2 graph coloring algorithm for distributed memory computers. *High Performance Computing and Communcations*, pages 796–806, 2005.

[BCS09] R. Braden, D. Clark, and S. Shenker. RFC 1633 — Integrated Services in the Internet architecture : an overview, June 2009.

[BDPV07] A. Bui, A.K. Datta, F. Petit, and V. Villain. Snap-stabilization and PIF in tree networks. *Distributed Computing*, 20(1) :3–19, 2007.

[Bea89] J. Beasly. An SST-based algorithm for the Steiner problem in graphs. *Networks*, 19, 1989.

[BEH+07] G. Di Battista, T. Erlebach, A. Hall, M. Patrignani, M. Pizzonia, and T. Schank. Computing the types of the relationships between autonomous systems. *IEEE/ACM Transactions on Networking*, 15(2) :267–280, April 2007.

[BEHV05] D. Barth, L. Echabbi, C. Hamlaoui, and S. Vial. An economic and algorithmic model for QoS provisioning BGP interdomain network. In *Proceedings of the EuroNGI Workshop on QoS and Traffic Control*, Paris, France, December 2005.

[Ber66] C. Berge. *The theory of graphs and its applications.* Wiley, 1966.

[Bet01] C. Bettstetter. Smooth is better than sharp : a random mobility model for simulation of wireless networks. In *Proceedings of the 4th ACM International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 19–27. ACM, 2001.

[BGRS10] J. Byrka, F. Grandoni, T. Rothvoß, and L. Sanita. An improved LP-based approximation for Steiner tree. In *Proceedings of ACM STOC*, 2010.

[BH04] F. Bai and A. Helmy. *Wireless ad hoc and sensor networks*, chapter A survey of mobility modeling and analysis in wireless ad hoc networks. Kluwer Academic Publishers, June 2004.

[BMM09] D. Barth, T. Mautor, and D. Villa Monteiro. Impact of alliances on end-to-end QoS satisfaction in an interdomain network. In *ICC*, pages 1–6, 2009.

[BT02] T. Bu and D. Towsley. On distinguishing between Internet power law topology generators. In *Proceedings of INFOCOM'02, 21st IEEE International Conference on Computer Communications*, volume 2, pages 638–647, New York, USA, June 2002.

[BZB+97] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin. RFC 2205 — Resource reSerVation Protocol (RSVP), version 1 functional specification, September 1997.

[C+06] F. Callegati et al. Research on optical core networks in the e-photon/one network of excellence. In *Proceedings of IEEE INFOCOM*, pages 1–5, 2006.

[cai] Caida AS links dataset. http ://www.caida.org/research/topology/.

[CAI02] CAIDA. IP paquet length distribution online, June 2002. `http ://www.caida.org/analysis/AIX/plen_hist`.

[CAR] CARRIOCAS. The CARRIOCAS project website. http ://www.carriocas.org.

[CGJ+04] H. Chang, R. Govindan, S. Jamin, S. Shenker, and W. Willinger. Towards capturing representative AS-level Internet topologies. *Computer Networks Journal*, 44(6) :737–755, April 2004.

[CH04] H. Castel and G. Hebuterne. Performance analysis of an optical MAN ring for asynchronous variable length packets. In *Proceedings of IEEE ICT*, Fortaleza, Brazil, August 2004.

[CN98a] S. Chen and K. Nahrstedt. An overview of quality-of-service routing for the next generation high-speed networks : Problems and solutions. *IEEE Networks*, 12(6) :64–79, December 1998.

[CN98b] S. Chen and K. Narhstedt. On finding multi-constrainted paths. In *Proceedings of ICC'98, IEEE International Conference on Communications*, volume 2, pages 874–879, Atlanta, USA, June 1998.

[CNSA10] D. Chiaroni, D. Neilson, C. Simonneau, and J-C. Antona. Novel optical packet switching nodes for metro and backbone networks. In *Proceedings of ONDM*, 2010.

[CO93] I. Cidon and Y. Ofek. MetaRing — a full duplex ring with fairness and spatial reuse. *IEEE Trans. on Comm.*, 41(1) :110–120, January 1993.

[CWK00] R. Cardwell, O. Wasem, and H. Kobrinski. WDM architecture and economics in metropolitan area. *SPIE Optical Network Mag.*, July 2000.

[Dav81] M. Davio. Kronecker product and shuffle algebra. *IEEE Trans on Comp.*, 20(2) :116–125, 1981.

[DIM97] S. Dolev, A. Israeli, and S. Moran. Uniform dynamic self-stabilizing leader election. *Distributed Algorithms*, pages 167–180, 1997.

[DKF+07] X. A. Dimitropoulos, D. V. Krioukov, M. Fomenkov, B. Huffaker, Y. Hyun, kc claffy, and G. F. Riley. AS relationships : inference and validation. *Computer Communication Review*, 37(1) :29–40, 2007.

[DM03] S. Dorogovtsev and J. Mendes. *From biological networks to the Internet and WWW.* Oxford University Press, January 2003.

[Doa96] M. Doar. A better model for generating test networks. In *Proceedings of GLOBECOM'96, IEEE Global Telecommunications Conference*, pages 86–93, London, Great Britain, 1996.

[DV07] P. Djukic and S. Valaee. Link scheduling for minimum delay in spatial re-use TDMA. In *Proc. of the 26th IEEE International Conference on Computer Communications (INFOCOM)*, pages 28–36, May 2007.

[EDF] EDF. Le rapport interne d'EDF pour le projet CARRIOCAS.

[EM04] Ed. E. Mannie. RFC 3945 — generalized multi-protocol label switching (GMPLS) architecture, October 2004.

[FAV08] A. Farrel, A. Ayyangar, and J.P. Vasseur. RFC 6051 — inter-domain MPLS and GMPLS traffic engineering — ressource reservation protocol — traffic engineering (RSVP-TE) extentions, February 2008.

[FK97] I. Foster and C. Kesselman. Globus : A metacomputing infrastructure toolkit. *Intl J. Supercomputer Applications*, 11(2) :115–128, 1997.

[FK01] I. Foster and C. Kesselman. The anatomy of the grid : Enabling scalable virtual organizations. *International Journal of High Performance Computing Applications*, 5(3) :200–202, 2001.

[FK02] I. Foster and C. Kesselman. *Computational Grids.* Morgan Kaufmann, 2002. The Grid : Blueprint for a New Computing Infrastructure.

[Fos05] I. Foster. Globus toolkit version 4 : Software for service-oriented systems. (LNCS 3779) :2–13, 2005.

[FPSS02] J. Feigenbaum, C. Papadimitriou, R. Sami, and S. Shenker. A BGP-based mechanism for lowest-cost routing. In *Proceedings of PODC'02, 21st Annual Symposium on Principles of Distributed Computing*, pages 173–182, Monterey, CA, USA, 2002.

[G+01] P. Gravey et al. Multiservice optical network : Main concepts and first achievements of the ROM Project. *IEEE Journal of Lightwave Technology*, 19 :23–31, January 2001.

[GALM08] Ph. Gill, M. F. Arlitt, Z. Li, and A. Mahanti. The flattening Internet topology : Natural evolution, unsightly barnacles or contrived collapse ? In *Proceedings of PAM*, pages 1–10, 2008.

[Gao01] L. Gao. On inferring autonomous system relationships in the Internet. *IEEE/ACM Transactions on Networking*, (9(6)) :73–745, 2001.

[GFJG01] Z. Ge, D. R. Figueiredo, S. Jaiswal, and L. Gao. On the hierarchical structure of the logical Internet graph. In *Proceedings of SPIE ITCOM'01, Conference on Multimedia Systems and Applications*, volume 4526, pages 208–222, Colorado, USA, July 2001.

[GGR01] L. Gao, T. G. Griffin, and J. Rexford. Inherently safe backup routing with BGP. In *Proceedings of INFOCOM'01, 20th Annual Joint Conference of the IEEE Computer and Communications Societies*, pages 547–556, April 2001.

[Gib85] A. Gibbons. *Algorithmic Graph Theory*. Cambridge University Press, 1985.

[GJ79] M. R. Garey and D. S. Johnson. *Computers and Intractability : A Guide to the Theory of NP-Completeness*. W. H. Freeman and company, 25th edition, 1979.

[GO97] R. Guerin and A. Orda. QoS-based routing in networks with inaccurate information : Theory and algorithms. In *INFOCOM'97 : Proceedings of 16th Annual Joint Conference of the IEEE Computer and Communications Societies*, pages 75–83, Japan, April 1997.

[GR00] L. Gao and J. Rexford. Stable Internet routing without global coordination. In *Proceedings of SIGMETRICS'00, ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems*, pages 307–317, Santa Clara, CA, USA, June 2000.

[GR03] O. Gerstel and H. Raza. On the synergy between electrical and photonic switching. *IEEE Communication Magazine*, (4) :98–104, April 2003.

[Grö00] J. Grönkvist. Assignment methods for spatial reuse TDMA. In *Proceedings of the 1st ACM International Symposium on Mobile ad hoc Networking & Computing (MobiHoc'00)*, pages 119–124, Piscataway, NJ, USA, 2000. IEEE Press.

[GSB+08] M. Gyarmati, U. Schilcher, G. Brandner, C. Bettstetter, Y. W. Chung, and Y. H. Kim. Impact of random mobility on the inhomogeneity of spatial distributions. In *Proceedings of IEEE GLOBECOM*, pages 404–408, 2008.

[GW02] L. Gao and F. Wang. The extent of AS path inflation by routing policies. In *Proceedings of GLOBECOM'02, IEEE Global Telecommunications Conference*, volume 3, pages 2180–2184, November 2002.

[HA00] D. K. Hunter and I. Andonovic. Approaches to optical Internet packet switching. *IEEE Communication Magazine*, (9) :116–122, September 2000.

[HGPC99] X. Hong, M. Gerla, G. Pei, and C.-C. Chiang. A group mobility model for ad hoc wireless networks. In *Proceedings of the 2nd ACM international workshop on modeling, analysis and simulation of wireless and mobile systems*, pages 53–60, 1999.

[Hil96] J. Hillston. *A Compositional Approach to Performance Modelling*. Cambridge University Press, 1996.

[HR92] F. Hwang and D. Richards. Steiner tree problems. *Networks*, 22 :55–89, 1992.

[HT00] J. Hillston and J. Tomasik. Amalgamation of transition sequences in the PEPA formalism. In *Proceedings of PAPM 2000, the 8th International Workshop on Process Algebra and Performance Modelling, ICALP Satellite Workshops*, pages 523–534, Géneve, Swiss, July 2000.

[HT04] T. Herman and S. Tixeuil. A distributed TDMA slot assignment algorithm for wireless sensor networks. *Algorithmic Aspects of Wireless Sensor Networks*, pages 45–58, 2004.

[HTWL96] T. Hsu, K. Tsai, D. Wang, and D. T. Lee. Steiner problems on directed acyclic graphs. In *COCOON'96 : Proceedings of the 2nd Annual International Conference on Computing and Combinatorics*, pages 21–30, 1996.

[IT] ITU-T. Terms and definitions related to quality of service and network performance including dependability. ITU-T Recommendation E.800.

[Jaf84] J.M. Jaffe. Algorithms for finding paths with multiple constraints. networks. 14(1) :95–116, 1984.

[JBRAS03] A. Jardosh, E. M. Belding-Royer, K. C. Almeroth, and S. Suri. Towards realistic mobility models for mobile ad hoc networks. In *Proceedings of the 9th annual international conference on Mobile computing and networking, MobiCom'03*, pages 217–229. ACM, 2003.

[JCJ00] C. Jin, Q. Chen, and S. Jamin. Inet : Internet topology generator. Technical report, Department of EECS, University of Michigan, 2000.

[JM96] D. B. Johnson and D. A. Maltz. Dynamic source routing in ad hoc wireless networks. In *Mobile Computing*, pages 153–181. Kluwer Academic Publishers, 1996.

[Joh73] D. S. Johnson. Approximation algorithms for combinatorial problems. In *Proceedings of the 5th Annual ACM Symposium on Theory of Computing (STOC'73)*, pages 38–49, Austin, TX, USA, 1973.

[Jun07] D. Jungnickel. *Graphs, Network and Algorithms.* Springer-Verlag, 3rd edition, 2007.

[KAP03] I. Kotuliak, T. Atmaca, and D. Popa. Performance issues in all-optical networks : Fixed and variable packet format. *Photonics in Switching*, November 2003. invited paper.

[Kar72] R. M. Karp. *Complexity of Computer Computation*, chapter Reductibility Among Combinatorials Problems, pages 85–106. 1972.

[KCC05] S. Kurkowski, T. Camp, and M. Colagrosso. Manet simulation studies : The incredibles. *ACM SIGMOBILE Mobile Computing and Communications Review*, 9 :50–61, 2005.

[KK01] T. Korkmaz and M. Krunz. A randomized algorithm for finding a path subject to multiple QoS requirements. *Computer Networks*, 36(2–3) :251–268, 2001.

[KM05] F.A. Kuipers and P. Van Mieghem. Conditions that impact the complexity of QoS routing. *IEEE/ACM Transactions on Networking*, 13(4) :717–730, 2005.

[KMPS05] V. S. Kumar, M. V. Marathe, S. Parthasarathy, and A. Srinivasan. Algorithmic aspects of capacity in wireless networks. In *Proc. of the ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS)*, 2005.

[KN05] M. Kodialam and T. Nandagopal. Characterizing the capacity region in multi-radio multi-channel wireless mesh networks. In *Proc. of the 11th ACM Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2005.

[KR03] J.F. Kurose and K.W. Ross. *Computer Networking : a Top-Down Approach Featuring the Internet.* Addison Wesley, 2nd edition, 2003.

[KT06] J. Kleinberg and E. Tardos. *Algorithm Design.* Pearson Education, 2006.

[KUHN03] A. Kanzaki, T. Uemukai, T. Hara, and S. Nishio. Dynamic TDMA slot assignment in ad hoc networks. In *Proc. of the 17th IEEE International Conference on Advanced Information Networking and Applications (AINA)*, 2003.

[Mat62] K. Matthes. Zür Theorie der Bedienungprozesse. In *The 3rd Prague Conference Inf. Th*, pages 513–528, 1962.

[MLMB01] A. Medina, A. Lakhina, I. Matta, and J. Byers. BRITE : An approach to universal topology generation. In *Proceedings of MASCOTS'01, 9th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*, Cincinnati, OH, USA, August 2001.

[MR05] J. Mihelic and B. Robic. Solving the k-center problem effciently with a dominating set algorithm. *J. of Comp. and Info. Tech. (CIT)*, 13(3) :225–233, 2005.

[MS97] Q. Ma and P. Steenkiste. Quality of service routing for traffic with performance guarantees. In *Proceedings of IFIP'97, 1st IEEE/IFIP International Workshop on Quality of Service*, pages 115–126, New York, NY, USA, May 1997.

[MT90] S. Martello and P. Toth. *Knapsack problems : algorithms and computer implementations*. John Wiley & Sons, Inc., 1990.

[MYC+00] J. Myoungki, X. Yijun, H.C. Cankaya, M. Vandenhoute, and C. Qiao. Effcient multicast schemes for optical burst-switched WDM networks. In *Proc. of IEEE ICC*, 2000.

[MZQ98] R. Malli, X. Zhang, and C. Qiao. Benefits of multicasting in all-optical networks. In *Proc. of SPIE, All Optical Networking*, pages 209–220, November 1998.

[NS07] NS. Network simulator [Online]. *Available : http ://ns-nam.isi.edu/nsnam/index.php/Main_ Page*, March 2007.

[O'M01] M. O'Mahony. The application of optical packet switching in future communications networks. *IEEE Comm. Mag.*, pages 125–128, March 2001.

[OPW+08] R. V. Oliveira, D. Pei, W. Willinger, B. Zhang, and L. Zhang. In search of the elusive ground truth : the Internet's AS-level connectivity structure. In *Proceedings of SIGMETRICS'08*, pages 217–228, 2008.

[Pa03] D. Popa and al. Evaluating the performance of an all-optical metro ring network : Fixed-length packet-switching versus variable-length packet-switching technology. In *Proceedings of HET-NETs*, Ilkley, UK, July 2003.

[PA06] D. Popa and T. Atmaca. A distributed fairness algorithm for bus-based metropolitan optical network. In *Proceedings of IEEE IPCCC*, Phoenix, USA, April 2006.

[PF91] B. Plateau and J.-M. Fourneau. Methodology for solving Markov models of parallel systems. *Journal of Parallel and Distributed Computing*, (12) :370–387, 1991.

[Pla84] B. Plateau. *De l'evaluation du parallelisme et de sa synchronisation.* Thèse d'état, Université de Paris-Sud, Centre d'Orsay, November 1984.

[PS85] F. P. Preparata and M. I. Shamos. *Computational Geometry.* Springer, Berlin Heidelberg, 1985.

[PTBV10] S. Pomportes, J. Tomasik, A. Busson, and V. Vèque. Self-stabilizing algorithm of two-hop conflict resolution. In *Proceedings of the 12th International Symposium on Stabilization, Safety, and Security of Distributed Systems (SSS 2010)*, pages 288–302, September 2010. LNCS 6366.

[PTBV11] S. Pomportes, J. Tomasik, A. Busson, and V. Vèque. Resource allocation in ad hoc networks with two-hop interference resolution. In *Proceedings of 54th IEEE Global Communications Conference (IEEE GLOBECOM)*, December 2011.

[PTV10] S. Pomportes, J. Tomasik, and V. Vèque. Ad hoc network in a disaster area : a composite mobility model and its evaluation. In *Proceedings of the 2010 International Conferences on Advanced Technologies for Communications (ATC 2010)*, Ho Chi Minh City, Vietnam, October 2010.

[PTV11] S. Pomportes, J. Tomasik, and V. Vèque. A composite mobility model for ad hoc networks in disaster areas. *REV Journal on Electronics and Communications*, 1(1) :62–68, January–March 2011.

[PTWB11] D. Poulain, J. Tomasik, M.-A. Weisser, and D. Barth. Optimal receiver cost and wavelength number minimization in all-optical ring networks. In *Proceedings of 11th International Conference on Telecommunications (ConTEL 2011)*, pages 411–417, June 2011.

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

[QT99] F. Quessette and J. Tomasik. Functional transitions in Stochastic Automata Networks. *Archives of Theoretical and Applied Computer Science*, 11(1) :69–85, 1999.

[Que94] F. Quessette. *"De nouvelles méthodes de résolution pour l'analyse quantitative des systèmes paralleles et des protocoles"*. PhD thesis, Université Paris-Sud, Orsay, November 1994.

[RB96] S. Robert and J.-Y. Le Boudec. On a Markov modulated chain exhibiting self-similarities over finite timescale. *Performance Evaluation*, 27&28, 1996.

[RCT+11a] V. Reinhard, J. Cohen, J. Tomasik, D. Barth, and M-A. Weisser. Optimal configuration of an optical network providing predefined multicast transmissions. *Computer Networks (Elsevier)*, 2011. accepté.

[RCT+11b] V. Reinhard, J. Cohen, J. Tomasik, D. Barth, and M-A. Weisser. Performance improvement of an optical network providing services based on multicast. In *Proceedings of ISCIS*, London, UK, Septembre 2011.

[RLH06] Y. Rekhter, T. Li, and S. Hares. RFC 4271 — a Border Gateway Protocol 4 (BGP-4), 2006.

[RMSM01] E. M. Royer, P. M. Melliar-Smith, and L. E. Moser. An analysis of the optimum node density for ad hoc mobile networks. In *Proceedings of of INFOCOM'01 : the IEEE International Conference on Communications*, pages 857–861, 2001.

[RSH+08] I. Rhee, M. Shin, S. Hong, K. Lee, and S. Chong. On the Levy-Walk nature of human mobility. In *Proc. of INFOCOM : IEEE Conference on Computer Communications*, pages 924–932. IEEE, April 2008.

[RT08a] V. Reinhard and J. Tomasik. A centralised control mechanism for network resource allocation in grid applications. *International Journal of Web and Grid Services*, 4(4) :461–475, 2008.

[RT08b] V. Reinhard and J. Tomasik. A model of application-network interactions in a high-speed optical network. In *Proceedings of 2nd International Workshop on P2P, Parallel, Grid and Internet Computing, 3PGIC*, Barcelone, Espagne, 2008. 6 pages.

[RTBW09] V. Reinhard, J. Tomasik, D. Barth, and M-A. Weisser. Bandwith optimisation for multicast transmissions in virtual circuits networks. In *Proceedings of IFIP Networking 2009*, pages 859–870, L'Aix-la-Chapelle, Allemagne, May 2009.

[RV99] B. Raghavachari and J. Veerasamy. Approximation algorithms for mixed postman problem. *SIAM Journal on Discrete Mathematics*, 12(4) :425–433, 1999.

[RVC01] E. Rosen, A. Viswanathan, and R. Callon. RFC 3031 — MultiProtocol Label Switching architecture, January 2001.

[RWMX06] I. Rhee, A. Warrier, J. Min, and L. Xu. DRAND : distributed randomized TDMA scheduling for wireless ad-hoc networks. In *Proc. of the 7th ACM International Symposium on Mobile ad hoc Networking and Computing (MobiHoc)*, 2006.

[RZ00] G. Robins and A. Zelikovsky. Improved Steiner tree approximation in graphs. In *Proceedings of ACM-SIAM SODA*, 2000.

[Saa80] Y. Saad. Variation on Arnoldi's method for computing eigenelements of large unsymmetric matrices. *Linear Algebra and Its Applications*, (34), 1980.

[Saa81] Y. Saad. Krylov subspace methods for solving unsymmetric linear systems. *Mathematics in Computations*, (37), 1981.

[Saa91] Y. Saad. *Projection methods for the numerical solution of Markow chain models.* Marcel Drekker, New York, NY, USA, 1991.

[Saa95] Y. Saad. *Preconditioned Krylov subspace medhods for the numerical solution of Markov chains.* Kluwer International Publishers, Boston, MA, USA, 1995.

[SARK02] L. Subramanian, S. Agarwal, J. Rexford, and R. H. Katz. Characterizing the internet hierarchy from multiple vantage points. In *Proceeding of INFOCOM'02, 21st IEEE International Conference on Computer Communications*, volume 2, pages 618–627, New York, USA, June 2002.

[SBMIY07] F. Sivrikaya, C. Busch, M. Magdon-Ismail, and B. Yener. ASAND : asynchronous slot assignment and neighbor discovery protocol for wireless networks. pages 27–31, Washington DC, USA, 2007.

[SDD+01] N. L. Sauze, A. Dupas, E. Dotaro, L. Ciavaglia, M. Nizam, A. Ge, and L. Dembeck. A novel, low cost optical packet metropolitan ring architecture. In *Proc. of ECOC*, pages 66–67, September–October 2001.

[SF04] G. Siganos and M. Faloutsos. Analyzing BGP policies : Methodology and tool. In *Proceedings of INFOCOM'04, 23rd Annual Joint Conference of the IEEE Computer and Communications Societies*, volume 3, March 2004.

[SFK+10] S. Shakkottai, M. Fomenkov, R. Koga, D. Krioukov, and kc claffy. Evolution of the Internet AS-level ecosystem. *European Physical Journal B*, 74(2) :271–278, March 2010.

[SRV97] H. F. Salama, D. S. Reeves, and Y. Viniotis. Evaluation of multicast routing algorithms forreal-time communication on high-speed networks. *IEEE J. on Sel. Areas in Comm.*, 15(3), 1997.

[Sta04] W. Stallings. *Data and Computer Communications.* Pearson Prentice Hall, 7th edition, 2004.

[STT02] T. Salonidis, L. Tassiulas, and R. Tassiulas. Distributed on-line schedule adaptation for balanced slot allocation in wireless ad hoc networks. In *Proc. of IEEE International Workshop on Quality of Service (IWQoS)*, Montreal, Canada, 2002.

[TDG+01] H. Tangmunarunkit, J. Doyle, R. Govindan, W. Willinger, S. Jamin, and S. Shenker. Does AS size determine degree in AS topology ? *SIGCOMM Comput. Commun. Rev.*, 31(5) :7–8, 2001.

[TEA00] J. Tomasik, H. ElBiaze, and T. Atmaca. Performance of a multiservice switch using stochastic automata networks. In *Proceedings of Advanced Simulation Technologies Conference*, Washington D.C., USA, April 2000.

[TH00a] J. Tomasik and J. Hillston. Amalgamation of transition sequences in the pepa formalism. Technical Report EDI-INF-RR-0013, Laboratory for Foundations of Computer Science, The University of Edinburgh, March 2000.

[TH00b] J. Tomasik and J. Hillston. Transforming PEPA models to obtain product form bounds. Technical Report EDI-INF-RR-0009, Laboratory for Foundations of Computer Science, The University of Edinburgh, February 2000.

[TK09] J. Tomasik and I. Kotuliak. *Current Research Progress of Optical Networks*, chapter Markovian Analysis of a Synchronous Optical Packet Switch, pages 45–64. Springer, 2009.

[TKA03] J. Tomasik, I. Kotuliak, and T. Atmaca. Markovian performance analysis of a synchronous optical packet switch. In *Proceedings of IEEE/ACM MASCOTS 2003*, pages 254–257, October 2003.

[TM80] H. Takahashi and A. Matsuyama. An approximate solution for the Steiner problem in graphs. *Math. Jap.*, 24(6) :573–577, 1980.

[Tom98] J. Tomasik. *The Stochastic Automata Network method for modelling elements of computer networks.* Phd thesis, Politechnika Śląska (Polytechnique Silésienne), Gliwice, July 1998. en polonais.

[Tom99] J. Tomasik. An impact of source modulation on an autocorrelation function of a generated data stream. In *National Conference Computer Networks*, pages 357–374, Zakopane, Poland, 1999.

[Tom02] J. Tomasik. On Generalised Semi-Markov Processes, their insensitivity, and possible applications. *Archives of Theoretical and Applied Computer Science*, 14(1), 2002.

[TP04] J. Tomasik and D. Popa. Markovian model of an edge node in all-optical synchronous double bus networks. In *Proceedings of Computer and Communication Networks, CCN*, pages 25–30, Cambridge, MA, USA, November 2004.

[TP07] J. Tomasik and D. Popa. On Markov chain modelling of asynchronous optical CSMA/CA protocol. In *Proceedings of IEEE/ACM MASCOTS 2007*, pages 360–365, Istanbul, Turkey, October 2007.

[TRGW02] H. Tangmunarunkit, S. Shenker R. Govindan, S. Jamin, and W. Willinger. Network topology generators : Degree-based vs. structural. In *Proceedings of SIGCOMM'02, ACM annual conference of the Special Interest Group on Data Communication*, volume 32, pages 147–159, Pittsburgh, PA, USA, 2002. ACM Press.

[TTC99] L. Tancevski, L. Tamil, and F. Callegati. Nondegenerate buffers : an approach for building large optical memories. *IEEE Photonics Technology Letters*, 11(8) :1072–1074, August 1999.

[TW10a] J. Tomasik and M.-A. Weisser. aSHIIP : autonomous generator of random Internet-like topologies with inter-domain hierarchy. In *Proceedings of IEEE/ACM MASCOTS 2010*, Miami Beach, FL, USA, August 2010.

[TW10b] J. Tomasik and M.-A. Weisser. Internet topology on AS-level : model, generation methods and tool. In *Proceedings of IEEE IPCCC 2010*, Albuquerque, NM, USA, December 2010.

[Tél96] France Télécom. Passeport Frame Relay NNI, 1996. user guide.

[VAB+08] D. Verchère, O. Audouin, B. Berde, A. Chiosi, R. Douville, H. Pouylau, P. Primet, M. Pasin, S. Soudan, D. Barth, C. Caderé, V. Reinhard, and J. Tomasik. Automatic network services aligned with grid application requirements in CARRIOCAS project. In *Proceedings of GridNets'08*, pages 196–205, 2008.

[vD93] N.M. van Dijk. *Queueing Networks and Product Forms : a System Approach.* John Wiley & Sons, 1993.

[Ver07] D. Verchère. Orchestrating optimally IT and network resource allocations for stringent distributed applications over ultrahigh bit rate transmission networks. In *European Conference and Exhibition on Optical Communication*, Berlin, Germany, 2007.

[Wax88] B. M. Waxman. Routing of multipoint connections. *IEEE Journal on Selected Areas in Communications*, 6(9) :1617–1622, 1988.

[WC96] Z. Wang and J. Crowcroft. Quality of service routing for supporting multimedia applications. *IEEE Journal on Selected Areas in Communications*, 14 :1228–1234, 1996.

[Wei07] M-A. Weisser. *La qualité de service dans le réseau inter-domaine Internet : algorithmes et modélisation*. PhD thesis, Supélec & Université de Versailles Saint Quentin (UVSQ), December 2007.

[WRSK03] I. White, M. Rogge, K. Shrikhande, and L. Kazovsky. A summary of the HORNET project : a next-generation metropolitan area network. *IEEE JSAC*, 21(9) :1478–1494, November 2003.

[WT05] M-A. Weisser and J. Tomasik. A distributed algorithm for resources provisioning in networks. In *EuroNGI Workshop on QoS and Traffic Control*, Paris, France, December 2005.

[WT06a] M-A. Weisser and J. Tomasik. A distributed algorithm for inter-domain resources provisioning. In *Proceeding of Second EuroNGI Conference*, pages 9–16, Valencia, Spain, April 2006.

[WT06b] M-A. Weisser and J. Tomasik. Inferring inter-domain relationships in randomly generated network models. In *EuroNGI Workshop on New Trends in Modelling : Quantitative Methods and Measurements*, Torino, Italy, June 2006.

[WT07a] M-A. Weisser and J. Tomasik. Automatic induction of inter-domain hierarchy in randomly generated network topologies. In *Proceeding of Springsim 2007 : the 10th Communication and Networking Simulation Symposium (CNS'07)*, pages 77–84, Norfolk, Virginia, USA, March 2007.

[WT07b] M-A. Weisser and J. Tomasik. Innapproximation proofs for directed Steiner tree and network problems. Technical report, PRiSM Laboratory, Université de Versailles Saint Quentin (UVSQ), June 2007.

[WTB08] M-A. Weisser, J. Tomasik, and D. Barth. *Congestion Avoiding Mechanism Based on Inter-Domain Hierarchy*, pages 470–481. Springer, Singapore, May 2008. LNCS 4982.

[WTB10] M-A. Weisser, J. Tomasik, and D. Barth. *Intelligent Quality of Service Technologies and Network Management : Models for Enhancing Communication*, chapter Exploiting the inter-domain hierarchy for the QoS management. IGI Global, 2010.

[WWL+06] W. Wang, Y. Wang, X. Y. Li, W. Z.Song, and O. Frieder. Efficient interference-aware TDMA link scheduling for static wireless networks. In *Proc. of the 12th ACM Annual International Conference on Mobile Computing and Networking (MobiCom)*, 2006.

[XLWN02] L. Xiao, K.-S. Lui, J. Wang, and K. Nahrsted. QoS extension to BGP. In *ICNP'02 : Proceedings of the 10th IEEE International Conference on Network Protocols*, pages 100–109, Paris, France, 2002.

[YDM00] S. Yao, S. Dixit, and B. Mukherjee. Advanced in photonic packet switching : an overview. *IEEE Comm. Mag.*, vol. 38, pp. 84-94, 2000.

[ZCB96] E. W. Zegura, K. L. Calvert, and S. Bhattacharjee. How to model an internetwork. In *Proceedings of INFOCOM'96, 15th IEEE International Conference on Computer Communications*, volume 2, pages 594–602, San Francisco, USA, March 1996.