UFR ESSTIN
École Doctorale IAEM Lorraine
DFD Automatique

Thèse
présentée pour l'obtention du
Doctorat de l'Université Henri Poincaré, Nancy 1
par
OLIVIER FERVEUR

Optimisation des architectures IP/MPLS de transport mutualisé

Composition du jury

Rapporteurs :	Pr. Didier Georges, INP Grenoble
	Pr. Zoubir Mammeri, IRIT
Examinateurs :	Pr. Hugues Mounier, IEF
	Mr. René Kopp, tuteur en entreprise, TDF
	Pr. Francis Lepage, directeur de thèse, CRAN
	Mr. Eric Gnaedinger, MCF, encadrant de thèse, CRAN
	Pr. Gilles Millerioux, CRAN

Thèse effectuée en partenariat avec :

SUPERVISORS :
Francis Lepage
Eric Gnaedinger
René Kopp

LOCATION :
Nancy

TIME FRAME :
November 2009

ABSTRACT

The IP technology was classicly used only in Internet's context. But since few years ago, this standard was re-use for new need like telephony and TV broadcast. This multi-use has as a consequence to screate a need of mutualised the core network. Today, the functionality provided by MPLS allow virtualisalisation of network infrastructure on a single IP core, leading to more complex resource management. In this thesis, we study the possibility of current management of the shared network and we propose the addition of a management mechanism called TDCN. We will demonstrate the many possibilities offered by this new system to provide an optimized core network use.

RÉSUMÉ

La technologie IP autrefois utilisée uniquement dans le cadre d'Internet est de nos jours le support de nombreuses autres applications dé-corrélées telle que la téléphonie ou encore la télévision. Cette multi-utilisation du standard a entraîné une volonté de mutualisation au sein du coeur de réseau. Aujourd'hui, les fonctionnalités apportées par MPLS permettent une virtualisalisation des infrastructures de réseau IP sur un unique coeur , entraînant une complexification de la gestion de la ressource. Dans cette thèse, nous étudierons les possibilités actuelles de gestion de ce réseau mutualisé et nous proposerons l'adjonction d'un mécanisme de gestion appelé TDCN. Nous démontrerons les nombreuses possibilités offertes par ce nouveau système pour ainsi optimiser l'utilisation du réseau.

«La faculté de citer est un substitut commode à l'intelligence»

— Sommerset Maugham

REMERCIEMENTS

Tous d'abord, je remercie tout particulièrement Eric Gnaedinger qui m'a incité à réaliser cette thèse et pour avoir encadré mon travail dès le D.E.A et m'avoir soutenu pendant toutes ces années.

Je tiens ensuite à remercier René Kopp, responsable du service ARC pour m'avoir proposé ce sujet dans le cadre d'une collaboration entre le CRAN et TDF. Je lui suis redevable de m'avoir accueilli dans son équipe où j'ai pu travailler sur cette thèse et sur bien d'autres sujets passionnants.

Je remercie également pour ses conseils avisés le Professeur Francis Lepage qui m'a fait l'honneur d'être mon directeur de thèse. Je conserverai d'excellents souvenirs de notre présentation au comité d'expert réseaux du CNRS.

Je suis également très reconnaissant à mon collègue Sébastien Lourdez pour son travail de relecture et plus généralement à l'ensemble de l'équipe TDF de Metz qui m'a supporté tout au long de ces années.

Je remercie particulièrement le Professeur Gilles Millerioux pour ses suggestions d'études dans le domaine les diagrammes de phases sans lesquelles je n'aurais pu aboutir.

Je suis très reconnaissant aux Professeurs Didier Georges et Zoubir Mammeri pour avoir accepté d'être les rapporteurs de mes travaux ainsi qu'au Professeur Hugues Mounier pour sa participation au jury en qualité d'examinateur.

Merci enfin à toute ma famille, en particulier Caroline, de m'avoir supporté et aidé.

TABLE DES MATIÈRES

TABLE DES FIGURES

LISTE DES TABLEAUX

ACRONYMS

DRY	Don't Repeat Yourself
API	Application Programming Interface
UML	Unified Modeling Language
TDCN	Tunnel Differentiation by Congestion Notification
ECN	Explicit Congestion Notification
HLA	High Level Autorization
LLA	Low Level Autorization
CLA	Current Level Autorization
RIO	RED In-profile Out-profile
RED	Random Early Detection
MPLS	MultiProtocol Label Switching
LDP	Label Distribution Protocol
RSVP	ReSerVation Protocol
PPVPN	Provider Provisioned Virtual Private Networks
SNMP	Simple Network Management Protocol
TE	Traffic Engineering
VPLS	Virtual Private LAN Service
IGP	Interior Gateway Protocol
TMS	Transport Multi-Services
TNT	Télévision Numérique Terrestre
RTT	Round Trip Time
DAR	Dynamique Acces Regulator

PRÉFACE

Cette thèse a été réalisée dans le cadre d'un partenariat entre le centre d'étude de TDF situé à Metz et le laboratoire CRAN (Centre en Automatique de Nancy - UMR 7039), unité mixte de recherche commune à l'Université Henri Poincaré, Nancy 1, à l'Institut National Polytechnique de Lorraine et au CNRS.

LE LABORATOIRE DU CRAN

Cette collaboration s'inscrit dans la thématique de recherche Systèmes contrôlés en réseaux du CRAN, dont le but est de proposer des modèles et des méthodes permettant de concevoir des applications industrielles distribuées sur un réseau de communication.

LA SOCIÉTÉ TDF

Partenaire des télévisions, radios, opérateurs de télécommunications et collectivités locales, TDF est un opérateur et un prestataire de services de référence dans les domaines de l'audiovisuel, de la téléphonie mobile et du haut débit.

Tournage vidéo, diffusion analogique et numérique de la télévision et de la radio, déploiement, maintenance et gestion de réseaux de télécommunications ; les services de TDF s'appuient sur une expertise reconnue, un parc hertzien de plus de 7 500 sites, une proximité des équipes et un service client de qualité.

Tourné vers l'avenir, le Groupe s'affirme au plan européen comme un acteur dynamique de la convergence entre audiovisuel et télécommunications (DVB-H), mais aussi comme partenaire majeur de l'aménagement numérique des territoires (TNT, WiMAX). Depuis 2002, TDF est une entreprise certifiée ISO 9001 version 2000 pour l'ensemble de ses services.

INTRODUCTION

Avec l'accélération de l'utilisation d'Internet, on assiste à une convergence des secteurs de la téléphonie fixe, mobile, des données et du monde audiovisuel. Les réseaux supportant ces services sont historiquement séparés et leurs interactions sont possibles par l'intermédiaire de passerelles centrales. Cependant, une mutualisation de l'infrastructure physique est actuellement mise en oeuvre tout en conservant une séparation protocolaire. Ce nouveau modèle entraîne une complexification de la gestion des coeurs de réseaux. En effet, ceux-ci sont soumis à des contraintes quasi contradictoires de disponibilité, de qualité ou de flexibilité.

La problématique réside alors dans l'optimisation des architectures. Ce problème peut ainsi être traité à la fois par une gestion des topologies adossée à une gestion protocolaire spécifique aux divers besoins. Les travaux de recherche se focalisant sur l'étude de la topologie tendent à améliorer la disponibilité, à contrario les études protocolaires sont axées sur la qualité de service. Cependant, ces approches ne prennent pas vraiment en considération la flexibilité au égard de la diversité des flux transportés (Audio/visuel, Internet, Téléphonie, Données privées).

Dans ce mémoire de thèse, nous nous proposons de réconcilier ces besoins contradictoires et en complément des approches classiques, nous développerons un nouveau mécanisme de régulation protocolaire offrant de nouvelles capacités de flexibilité.

Le premier chapitre décrit plus précisément le contexte de notre étude. Après une description de l'architecture du réseau mutualisé (TMS : Transport Multi-Services), nous présenterons la problématique de la gestion interne du "Core Network".

Le chapitre suivant dresse un état de l'art des différents mécanismes de la régulation de trafic. Nous séparerons les approches orientées topologies (Provisioning), des approches protocolaires (ECN,XCP,...).

Le chapitre trois analyse ces diverses propositions et met en évidence leurs manques dans le cadre d'infrastructures mutualisées. Cette étude nous amènera ainsi à définir une approche novatrice fournissant les mécanismes manquant à une régulation optimale.

Le chapitre quatre s'attache à valider notre démarche. Pour ce faire nous avons effectué une modélisation formelle de cette approche.

L'intérêt de cette proposition est démontré dans le cadre d'une étude pratique, complétée par une étude d'optimisation du système.

Le dernier chapitre repose sur la réalisation d'une maquette. Celle-ci nous a permis d'une part de contrôler les interactions protocolaires, et d'autre part de mesurer les performances de notre nouveau système de régulation.

1

CONTEXTE

1.1 PRÉSENTATION DU PROJET TMS

Afin de réaliser mes travaux de recherche, j'ai été intégré à l'équipe de conception du projet TMS (Transport Multi-Service). Ce projet débuté en 2006 a pour objectif de réaliser un réseau de transport national pour l'ensemble des applications métiers de l'entreprise. Ce chapitre s'attarde à la présentation du projet TMS qui est le support principal de mes recherches.

1.1.1 *Descriptif général*

TDF met en place un nouveau réseau de transport IP/ Ethernet utilisant le protocole "Multiprotocol Label Switching" (MPLS) orienté audiovisuel. Les orientations techniques de ce réseau reposent sur une architecture moderne et résistante, tirant partie des développements technologiques du monde des télécommunications :

- une transmission en numérique, à haut débit ;
- un coeur de réseau en fibres optiques, avec une architecture en boucles ;
- un réseau d'accès aux extrémités en faisceaux hertziens numériques haut débit.
- un système de routage permettant la souplesse de configuration et l'introduction de nouveaux services.

1.1.2 *Objectifs*

Aujourd'hui, chaque demande de transport ou de raccordement est traitée au cas par cas, au moyen de liaisons individuelles achetées pour la plupart à l'opérateur FT ou à sa filiale Globecast. Les offres ne permettent pas de faire bénéficier nos clients d'un effet de mutualisation des moyens.
TMS permettra de partager techniquement et commercialement un réseau de télécommunication intégré et opéré par TDF entre les différents clients, tout en préservant une confidentialité absolue entre les

flux de ses clients.
Cette mutualisation se traduira par des offres commercialement plus attractives, plus souples pour les clients de TDF.
Par l'utilisation de ce réseau numérique, les clients pourront demander de nouveaux services, que TDF ne peut offrir actuellement.

1.1.3 *Une architecture Nationale*

Le réseau est composé d'un ensemble de boucles optiques étendues sur le territoire français. En plus de ces boucles, sont rajoutés des pendulaires en faisceaux hertziens. Au total, on dénombre environ 50 sites reliés en optique et plus de 150 en hertzien.

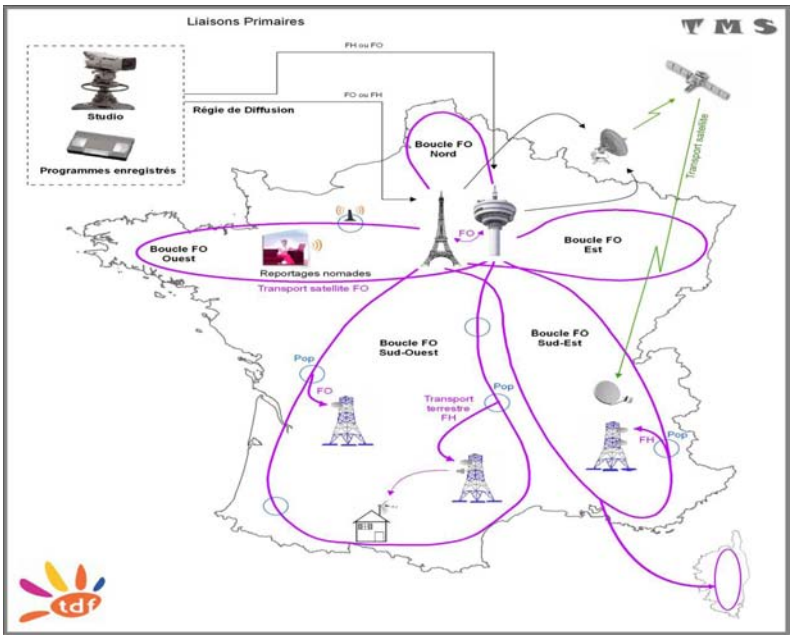

FIGURE 1: Principe du réseau TMS

1.1.4 *L'architecture protocolaire*

L'architecture protocolaire du réseau est conçue pour satisfaire des critères de résilience, de sécurité et de qualité. Tout d'abord, son

architecture en boucles permet de supporter la défaillance d'une liaison WAN.

L'ensemble du réseau utilise un adressage inclus dans une aire OSPF. Cette configuration facilite la création de tunnels MPLS. En effet, les protocoles de signalisation LDP utilisent les SPF pour la création automatique de tunnels.

Afin d'assurer le cloisonnement du réseau, aucun peering BGP n'est effectué avec un autre opérateur. Le coeur de réseau IP est uniquement utilisé pour faire transiter des flux de contrôle et de métrologie.

L'ensemble des inter-connexions et services apportés aux entreprises est encapsulé dans des messages MPLS de telle manière à assurer l'intégrité du coeur et l'indépendance entre les divers services.

1.1.5 *La technologie MPLS dans TMS*

MPLS est un mécanisme de transport de données, opérant sur la couche liaison, en dessous de protocoles comme IP. Il a été conçu pour fournir un service unifié de transport de données pour les clients sur la base de commutation de paquets ou commutation de circuits. MPLS peut être utilisé pour transporter différents types de trafic, par exemple de la voix ou des paquets IP.
MPLS fonctionne par commutation de labels. De façon manuelle ou automatique, l'administrateur du réseau MPLS établit un ou plusieurs chemins. On fait la différence en MPLS entre les routeurs d'entrée, de transit, et de sortie. Un chemin MPLS étant toujours unidirectionnel, le routeur d'entrée diffère du routeur de sortie.
Le routeur d'entrée a pour rôle d'encapsuler pour la première fois le trafic reçu sur ses interfaces « clients ». Il applique un label au paquet reçu et l'envoie vers une de ses interfaces sortantes.

La technologie MPLS possède un avantage de taille dans un réseau multiservices, elle permet de réaliser de l'ingénierie de trafic, c'est-à-dire de garantir la qualité de service et d'optimiser les ressources réseau. En effet, le chemin réseau pouvant être explicitement défini, MPLS possède les atouts pour réaliser une ingénierie de trafic fine.

De plus, aucune information des paquets clients n'est nécessaire pour effectuer ce routage. On peut donc transporter des flux non IP ou d'autres technologies. Il est ainsi possible de transporter des flux ATM par exemple.

On peut également adjoindre un véritable réseau en overlay en émulant les fonctionnements de switchs, de routeurs ou de méca-

nismes divers (MC-LAG[7], Pseudo-wire redundancy[1],...) relié par l'intermédiaire des liens MPLS.

MPLS garantit également une très grande disponibilité avec des protocoles tels que le Fast Reroute, qui permet de rétablir un lien rompu en moins de cinquante millisecondes.

1.1.6 *Les services TMS*

TMS a pour vocation le transport de divers services. Tous ces services sont réalisés à l'aide de PPVPN [3]. La topologie et les protocoles mis en place sont choisis sur mesure en fonction du service et de ses contraintes. Ainsi, un premier choix est effectué entre une virtualisation de niveau 2 (type VPLS [24]) ou niveau 3 (VPRN [17]), puis des compléments sont ajoutés pour respecter des contraintes de transport (H-VPLS [2],...).

Voici une série de services types pouvant circuler sur le réseau TMS. Cette liste est non-exhaustive et permet simplement de se rendre compte de la diversité des flux et des contraintes.

La diffusion de la TNT vers les émetteurs

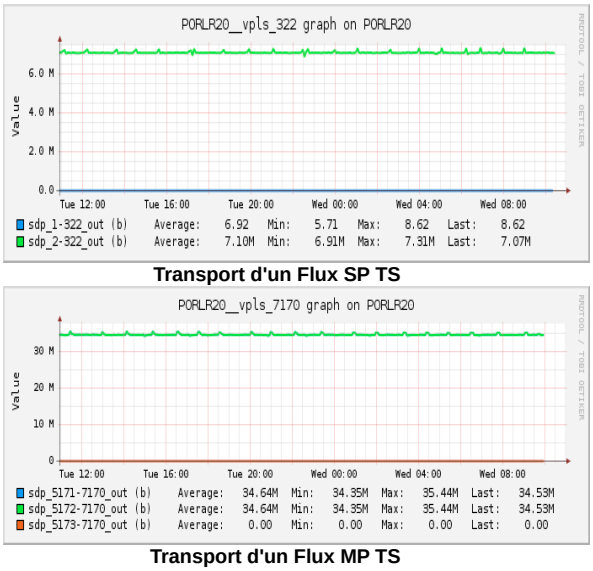

FIGURE 2: Evolution au court du temps d'un trafic de diffusion Vidéo

Du transport de la tête de réseau vers l'ensemble des émetteurs, ce service demande une contrainte de qualité forte qui s'exprime par un faible taux de perte paquets et une gigue minimale. Ce service se traduit par une diffusion multicast à fort débit. Le transport concerne des émetteurs possèdant une forte densité de population. Les éventuels re-routages se doivent d'être réalisés dans des délais inférieurs à la seconde.

Le transport sécurisé de multiplex audio-visuel

Ce service a pour mission le transport de flux audio-visuel en temps réel. La principale contrainte technique est l'impossibilité de l'interrompre à n'importe quel moment de la journée. Ce type de transport est notamment utilisé pour le transport vers les stations montantes satellites de diffusion vidéo. On retrouve les même contraintes que dans les services précédents avec des débits encore plus importants puisque transportant des bouquets complets de chaînes télévisuelles (Fig. 2). On peut également retrouver des encodeurs dans la chaîne si le transport concerne des chaînes analogiques.

Le transport d'agrégats de trafic Wimax

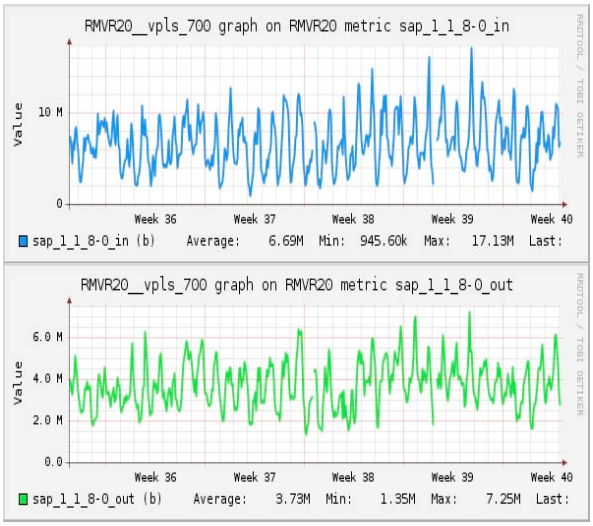

FIGURE 3: Evolution au court du temps d'un flux Internet agrégé

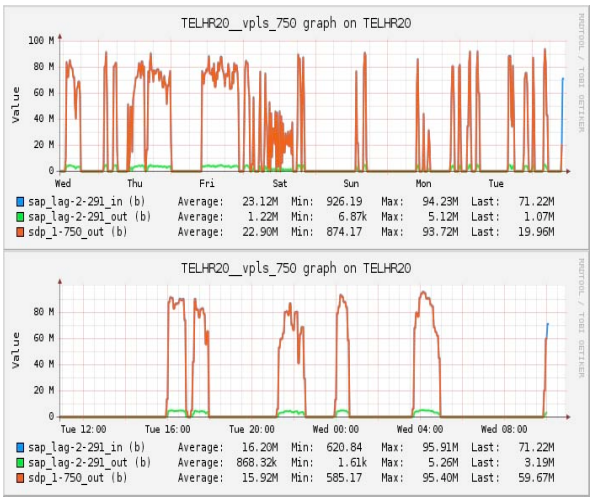

FIGURE 4: Evolution au court du temps d'un transport sporadique

Le réseau TMS est également le support de transport national pour des flux Internet, TDF ayant à sa charge un certain nombre de plaques régionales Wimax. L'ensemble des flux est agrégé avant d'être transporté de manière globale vers les différentes plate-formes des fournisseurs d'accès Internet. On y retrouve les profils cycliques classiques (Fig. 3) de l'utilisation d'Internet dus à la différence d'utilisation entre le jour et la nuit, le week-end et la semaine.

L'interconnexion métier

Un autre exemple remarquable concerne le service à l'entreprise. En effet, le besoin d'une grande quantité de bande passante sporadique peut apparaître(Fig. 4). Le backup de bases de données ou le transport de séquences audio-visuelles en vue de montage sont des exemples de ces cas d'utilisation.

1.1.7 *Les problématiques de TMS*

Les nombreuses particularités de ce réseau face aux contraintes habituellement exposées dans les réseaux IP nous amènent à nous interroger sur la pertinence des méthodes habituelles de gestion.

La mise en place d'infrastructures physiques, telles que les faisceaux hertziens et les fibres optiques, sont des postes de coûts importants pour le réseau. Aussi, une utilisation maximale de la bande passante est souhaitée et une gestion fine de l'ingénierie de trafic se doit d'être mise

en place. A contrario, les besoins sporadiques de bandes passantes élevées des clients sont en complète contradiction par rapport à une utilisation maximale du coeur de réseau. Les politiques classiques utilisées par les opérateurs traditionnels ne peuvent pas convenir sur ce réseau particulier du fait des hypothèses généralement formulées. En effet, on peut constater de nombreuses différences :

- Le réseau TMS dispose d'une majorité de flux RTP/UDP audio-visuels à fort débit.
- Etant donné sa position d'opérateur de gros, TMS ne dispose pas d'une base de clients "Best Effort à faible débit". La conséquence immédiate est qu'un provisionning statistique basé sur la théorie des grands nombres n'est pas adapté au système.
- Chaque client dispose de profils de trafics très différents voire changeants, ce qui, associés à leurs trafics élevés, remet également en cause une gestion de provisionning classique.

Finalement, on peut constater qu'une approche classique nous amène à refuser en accès de nombreux paquets qui finalement pourraient traverser sans encombre le réseau. Ce refus acceptable dans un réseau classique est plus problématique dans TMS.

En effet, de nombreux sites sont accessibles uniquement par faisceaux hertziens et il est alors beaucoup plus compliqué d'augmenter la bande passante. Un partage et une mise en concurrence des clients deviennent alors une quasi obligation.

La nécessité de ce partage est en plus renforcée par les évolutions des technologies des faisceaux. Les constructeurs conscients de ces limites tendent à réaliser un codage dynamique entraînant des bandes passantes qui varient dans le temps [19](en fonction de la météo notamment).

1.2 LE TRANSPORT MUTUALISÉ

Le chapitre précédent a mis en évidence un certain nombre de questionnements propres au réseau TMS. On peut finalement constater que ces interrogations sont propres à une famille de réseaux qui recherchent la fonction de transport mutualisé.

1.2.1 *Les opérateurs de Wholesale*

La convergence des réseaux et services engendre une mutualisation des infrastructures. Les différents services portés le plus souvent par

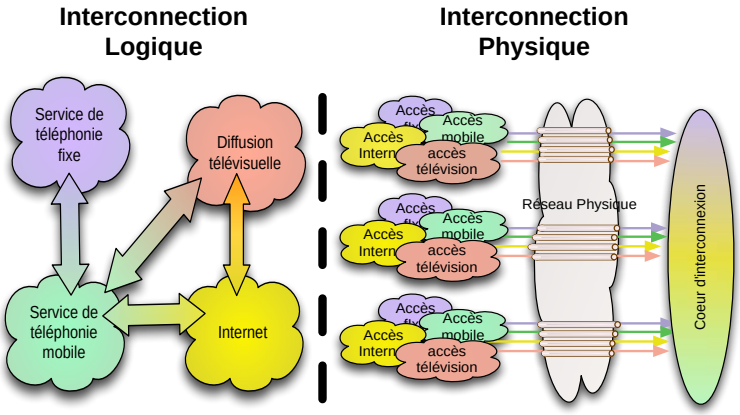

FIGURE 5: Positionnement d'un réseau de wholesale

des acteurs séparés regroupent leur infrastructure, soit en effectuant un rapprochement, soit en faisant appel à un opérateur mutualisant les ressources. Le dit opérateur est alors appelé "Opérateur de Wholesale". On retrouve alors les deux objectifs antagonistes présents dans le cadre de TMS :

- L'opérateur de "Wholesale" veut maximiser l'utilisation de son coeur pour profiter au mieux des bénéfices économiques apportés par la mutualisation.
- L'opérateur "virtuel" souhaite disposer d'un maximum de qualité, de resilience et de flexibilité pour d'une part se différencier de la concurrence et d'autre part fournir le meilleur et le plus attractif des services.

Dans ce cadre, la politique de qualité de service cible doit permettre de gérer les congestions dans le coeur du réseau, tout en minimisant les effets induits sur les extrémités. Ces opérateurs utilisent alors des technologies de qualité de service et des méthodes de gestion de bande passante ("Trafic ingenierie") pour effectuer ces optimisations.

1.2.2 *Exemple d'opérateur de wholesale : COLT*

Un exemple bien connu des réseaux de transport mutualisé est le réseau de la société COLT. Cette entreprise propose la mise à disposition de VPN pour les PME et les grandes entreprises européennes. Le réseau est composé principalement de liaisons optiques louées à différents opérateurs nationaux. Bien que positionné à une échelle

géographique différente du réseau TMS, la problématique reste similaire du fait du choix de la clientèle cible. On retrouve la nécessité de pouvoir cloisonner les divers clients tout en recherchant la meilleure optimisation pour le coeur de réseau.

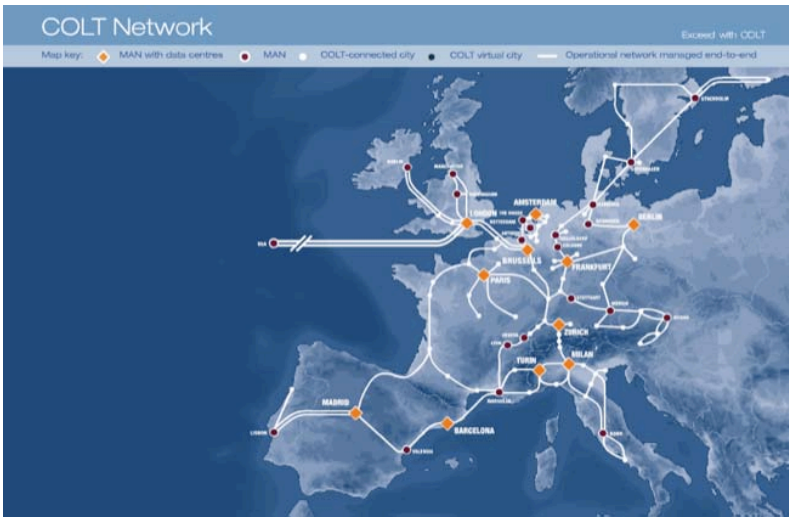

FIGURE 6: Réseau européen de la société COLT

1.2.3 *Hypothèses et bornes de l'étude*

Dans ce contexte, nous nous focaliserons sur ces méthodes de gestion. Nous concentrerons donc nos travaux aux contextes de "wholesale". Nous conserverons donc un certain nombres d'hypothèses :

- tous les flux des opérateurs virtuels seront encapsulés dans des PPVPN,
- l'ensemble des PPVPN sera réalisé à l'aide la technologie MPLS,
- toutes nos modifications seront effectuées dans le réseau à la charge de l'opérateur de Wholesale.

Nous étudierons les méthodes que nous qualifierons de "macroscopique" utilisées par les opérateurs pour dimensionner le débit des liaisons de manière globale.

Puis, nous étudierons les approches qualifiées de "microscopique" régulant la bande passante lors de congestion. Ces études nous permettrons de comprendre comment optimiser ces approches et adjoindre

de nouveaux mécanismes apportant une plus grande flexibilité aux opérateurs.

Derrière le concept de flexibilité, nous positionnerons principalement la capacité de pouvoir fournir de la bande passante opportuniste à un client (mutualisation de la bande passante).

2

ETUDE DE L'EXISTANT

2.1 PLANIFICATION RÉSEAUX

Afin de concevoir les mécanismes qui optimiseront au mieux le coeur de réseau, il faut comprendre comment est effectué le dimensionnement des liaisons et comment on gère cette problématique au jour le jour.

2.1.1 *Principe Général*

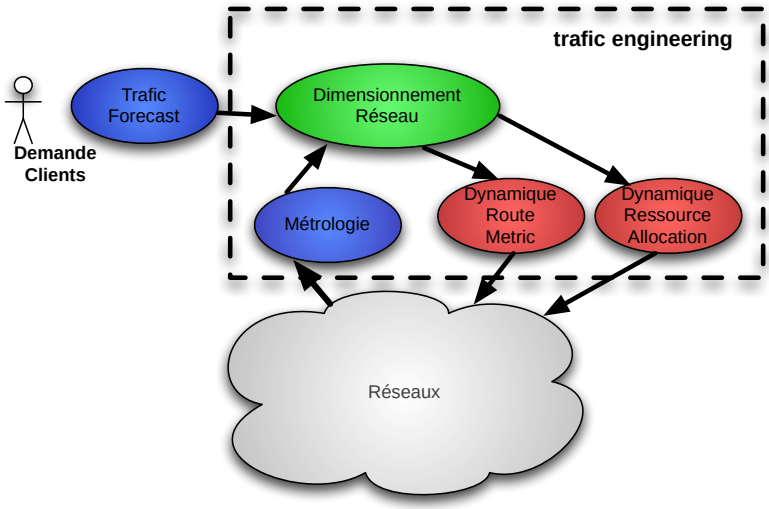

FIGURE 7: Principe du trafic engineering

Le principe du provisionning réseau repose sur une étude des trafics existants. Il existe de nombreuses méthodes systématiques de traitement pour cette problématique ([38] et [12]). De manière générale, on retrouve un certain nombre de concepts caractéristiques afin de répondre à une question simple : "Quel débit pour quelle liaison ?".

Ces concepts et les diverses propositions seront détaillés dans les paragraphes suivants.

2.1.2 *La matrice de trafic*

Il s'agit le plus souvent de la base des calculs. Cette matrice tente de déterminer quels routeurs communiquent avec quels autres routeurs et à quels débits. Dans un réseau IP standard (sans mécanisme de type MPLS), tous les trafics d'une source allant vers une destination passent par le même chemin à un moment donné.

A partir de ce constat et de la matrice, on peut ainsi prédire les lieux sous dimensionnés, et prévoir des reroutages, ou encore l'accroissement de certaines liaisons.

La détermination de la matrice de trafic est un travail en soit. En effet, le trafic de chaque noeud évoluant au cours de la journée et pouvant varier également à plus long terme, la détermination d'une valeur représentative n'est pas aisée. La valeur utilisée résulte le plus souvent d'une étude statistique des trafics observés sur le réseau [32].

Cette détermination se base sur une métrologie SNMP (Simple Network Management Protocol) interrogeant les équipements toutes les 5 minutes. Les phénomènes de faible durée peuvent être ignorés [31]. Ainsi, des micro-congestions peuvent apparaître et être totalement ignorées dans l'établissement de la planification.

Les travaux de C. Fraleigh [16] sont un exemple de détermination de cette matrice. A noter que le choix de cette matrice peut fortement impacter les résultats. Ainsi, C. Fraleigh tente de minimiser ces micro-congestions par le biais d'une estimation des profils clients.

2.1.3 *Optimisation des routes*

Une fois la matrice de trafic déterminée, plusieurs actions sont possibles. Les chemins réseaux utilisés pour aller d'un point A au point B sont dépendants uniquement de l'IGP (Interior Gateway Protocol). Celui-ci va favoriser une route en fonction de ses métriques. On peut alors calculer ces métriques pour optimiser la répartition. Dans la majorité des cas, on recherche un partitionnement le plus homogène possible. P. Sousa [36] va jusqu'à proposer une automatisation de ce processus.

2.1.4 *Apport de MPLS dans l'optimisation*

MPLS est, dans le cadre de ces calculs de routes optimales, un nouvel atout. En effet, l'hypothèse définie précédemment "tout trafic allant d'un point A vers un point B passe à un instant t par un chemin unique" peut ainsi être supprimée. En effet, chaque tunnel MPLS peut même, si ses extrémités sont communes, utiliser des chemins différents. M. Banner [5] démontre de façon théorique les apports de cette capacité à utiliser plusieurs chemins. De son côté, S. Kohler[22] s'attache à montrer l'efficacité d'une approche mixte "Traffic engineering" et optimisation de l'IGP.

2.1.5 *Contrôle d'admission*

On retrouve aussi des approches par contrôle d'admission [8] qui proposent que chaque routeur d'extrémité évalue dynamiquement la bande passante disponible pour aller vers les autres routeurs d'extrémité. Cette capacité est estimée à l'aide des fonctions TE des protocoles de routage. A chaque nouvelle requête client, le routeur d'extrémité vérifie la bande passante disponible et établit le chemin grâce à un algorithme "Constraint Based Routing" [40]. Ces approches permettent d'arriver pas à pas à un réseau le plus chargé possible sans congestion.

2.1.6 *Notion de résilience*

Cependant, d'autres contraintes sont à prendre en compte. En effet, le réseau n'est pas exempt de pannes (qualifiées par A. Markopoulou [29]). La panne d'une liaison entraîne un délestage vers les autres chemins. Si une optimisation est effectuée au plus juste, la simple panne d'une liaison peut compromettre une partie entière du réseau.

Pour éviter ce type de problème, la prise en compte des effets des pannes a été ajoutée pour considérer le débit présent sur les liaisons et le plus souvent, en utilisant la capacité de MPLS à signaler le débit utilisé (débit estimé au sens défini par la matrice de trafic) sur chacune des liaisons. De même, on réserve a priori la bande passante sur un chemin de backup disjoint de telle manière à assurer la résilience à une panne. Cet ajout entraîne une sur-évaluation importante de la bande passante utilisée. M. Kodialam[21] propose une mutualisation de chemins de backup issus de tunnels différents afin de limiter cette sur-évaluation.

2.1.7 *Réserve sur la planification réseau*

Des limites apparaissent quant à l'utilisation de ces méthodes de planification. En effet, on observe une sur-estimation de la bande passante nécessaire par liaison ("Over-provisioning"), d'une part à cause des marges prises lors de l'établissement de la matrice de trafic, et d'autre part à cause de la gestion de la panne. L'ensemble de ces mesures oblige à provisionner de la bande passante qui ne sera quasi jamais utilisée.

De plus, même si les approches par contrôle d'admission vont permettre une meilleure utilisation du coeur, elles ne peuvent répondre aux changements brutaux de bande passante qui se produisent dans nos exemples de wholesaling. Elles ne pourront pas non plus répondre aux capacités de liaisons changeantes [19].

Finalement, les améliorations apportées par la planification réseau ne permettent pas d'utiliser le maximum des capacités du réseau. Ces solutions maintiennent continuellement une partie de la bande passante non utilisée. Par la suite, nous nous focaliserons sur une méthode pour pouvoir utiliser cette bande passante laissée pour compte.

2.2 EVOLUTIONS DES SYSTÈMES DE RÉGULATION

Dans la plupart des réseaux, le point de congestion des flux élastiques se trouve en accès de celui-ci. Ceci peut s'expliquer historiquement par la rupture technologie entre les liaisons d'accès et de coeur (Digital Subscriber Line DSL / Wavelength Division Multiplexing WDM). Dans le cas des réseaux de transport mutualisé, cette rupture n'existe pas et la limitation est donc faite par policing. Ce qui entraîne la présence d'un point de congestion logique (par configuration) en accès du réseau.

Utiliser la bande passante résiduelle à la planification revient donc à lever ces limitations logiques de policing en accès. Ce qui nous ramène à devoir gérer les congestions au coeur du réseau et plus particulièrement à étudier le partage de la bande passante pendant cette congestion.

De nombreuses études sur l'équité ont été menées dans le cadre du réseau Internet. Nous allons tout d'abord revenir sur ces approches pour comprendre leurs apports et leurs limitations afin de rechercher une solution appropriée.

2.2.1 *Les mécanismes de la régulation de trafic*

Historiquement, la régulation de la bande passante dans les réseaux IP est reportée aux extrémités. Dans la pratique, seules les connexions TCP présentes sur les extrémités vont s'auto-limiter pour permettre à l'ensemble du réseau de bien se comporter. Pour optimiser cette approche, de nombreuses modifications sont régulièrement apportées à cette auto-régulation.

Des mécanismes de coeur peuvent venir compléter cette régulation. Par exemple, le mécanisme RED [9] (Ramdom Early Detection) supprime des paquets aléatoirement dans la file d'attente d'un noeud si celle-ci dépasse un seuil de remplissage déterminé. En rejetant des paquets TCP, il force la régulation des connexions à la baisse.

D'autres mécanismes agissent plus subtilement, la configuration d'un buffer dans le coeur peut entraîner une augmentation de la latence et par conséquent, une augmentation du RTT pour TCP ; ce qui là encore impacte le débit des connexions.

L'approche Diffserv [6] est un mécanisme connu qui permet de modifier la gestion de la bande passante. Cette approche agit indirectement sur la régulation produite par TCP. En effet, la mise en place d'une gestion de priorité entraîne nécessairement un allongement des délais de transit dans les noeuds modifiant ainsi le RTT de TCP.

On retrouve plusieurs motivations à l'ajout de ces mécanismes. La première repose sur l'amélioration qualitative de la congestion. L'objectif étant alors, soit de minimiser les pertes, soit de diminuer les délais de transit.

La seconde motivation réside alors dans l'établissement d'un partage équitable entre les sessions [28]. L'équité pouvant être définie de différentes manières. P. Marbach propose par exemple un partage reposant sur l'utilisation d'une pondération pour favoriser certains flux.

2.2.2 *L'approche Système Service Allocation*

Une des évolutions les plus intéressantes réside probablement dans le système Service Allocation [13]. Il repose sur des colorieurs situés en accès du réseau qui séparent le trafic TCP en 3 couleurs (vert, jaune, rouge). Les paquets de couleur rouge ne sont généralement pas admis dans le réseau et ceux de couleurs verte et jaune sont transmis dans le coeur avec des marqueurs de qualité de services différents. Les différents flux sont transportés dans les routeurs de coeur de réseau

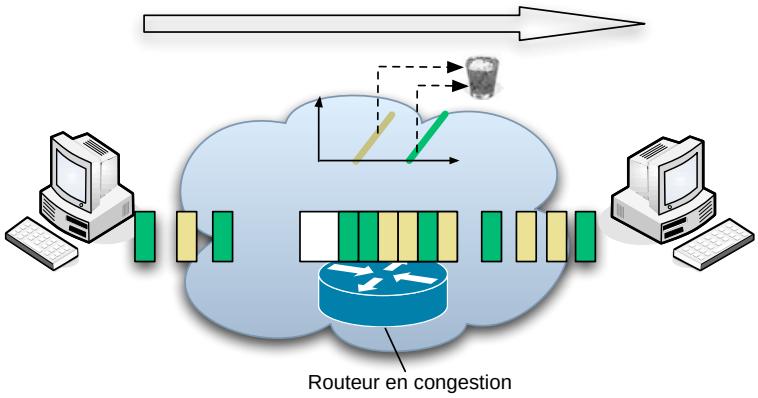

FIGURE 8: Principe du Système Service Allocation

par le biais d'un seul et même buffer. Cependant, la bufferisation est soumise à plusieurs systèmes RED. Chaque RED associé à une couleur dispose d'un paramètrage différent fixant son propre seuil de drop (on appel alors l'ensemble des mécanismes RIO pour "RED In Out"). Comparé aux systèmes à plusieurs buffers, chacun affecté à une qualité, l'intérêt de ce système réside dans le fait qu'il n'entraîne aucun désordonnancement.

Le but recherché est de fournir à un flux de données une qualité paramétrable par l'intermédiaire du colorieur d'accès. Toutefois de nombreuses études [14], [35], [26] ont démontré que ce système ne permet pas de garantir le respect d'une équité souhaitée. En effet, il est impossible de garantir un minimum de bande passante à chaque utilisateur. Si la première couleur est en saturation le système ne peut pas garantir ce minimum (nécessité d'une planification adapté).

Cette approche engendre également des difficultés dans le dimensionnement des buffers. Le remplissage du buffer d'émission d'une liaison en congestion est dépendant d'une part de la configuration du mécanisme RED et d'autre part du nombre de connexions TCP [37]. Cependant le nombre de connexions TCP pouvant varier fortement, les congestions apparaissant le plus souvent suite à des pannes de liaisons, il est impossible de prédire la nouvelle distribution de ces connexions d'autant plus que ce nombre reste lié au besoin de chaque client. On se retrouve alors avec un algorithme RED dont le paramètrage n'est plus adapté entraînant une baisse de la qualité de la liaison en terme de latence et de pertes.

En conclusion ce système reste difficilement applicable. Cependant, son concept apporte une vision intéressante car il ne propose pas de nouveaux mécanismes de coeur pour la gestion de la bande passante mais repose sur un assemblage de plusieurs mécanismes existants (RED + Marqueur + Colorieur).

2.2.3 *L'approche Système ECN*

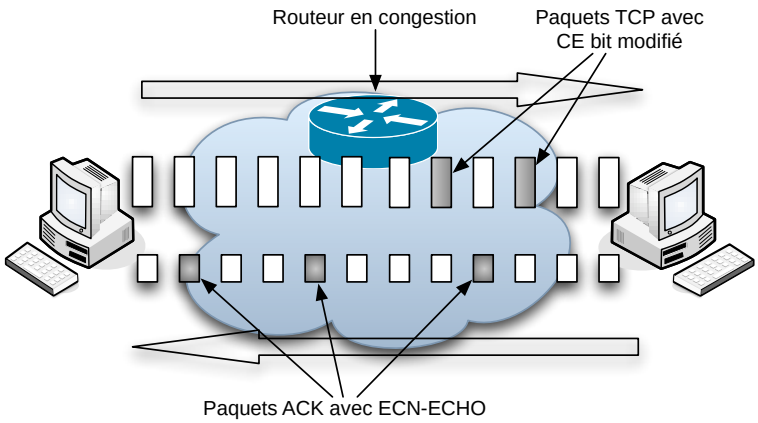

FIGURE 9: Principe du système ECN

L'approche ECN [33] va plus loin dans l'interaction des routeurs de coeur et des algorithmes d'accès. Cette approche, issue de mécanismes mise en oeuvre dans les technologies ATM puis Frame Relay, consiste à modifier le comportement de TCP par le biais du transport d'une information supplémentaire de congestion. Lors d'une congestion, les routeurs du coeur de réseau modifient à la volée le champs CE des paquets TCP. Le récepteur prend alors en compte cette information et la renvoie par l'intermédiaire de paquets ACK avec un champ ECN-ECHO. L'émetteur peut alors adapter son débit pour supprimer cette congestion sans perte de paquets.

Cette approche a donc conceptuellement un intérêt majeur : elle rapproche explicitement les mécanismes de régulation de coeur et d'accès.

Par contre, cette modification a plusieurs conséquences :

– elle oblige le routeur de coeur à modifier à la volée tous les paquets TCP, ce qui engendre des traitements supplémentaires du routeur pour gérer les paquets.
– la modification du paquet n'a d'intérêt que si les deux équipements d'extrémités utilisent cette information et modifient leurs comportements en conséquence. Afin d'obtenir l'équité recherchée l'ensemble des équipements terminaux doivent agir de manière homogène.

2.2.4 *eXplicit Control Protocol (XCP)*

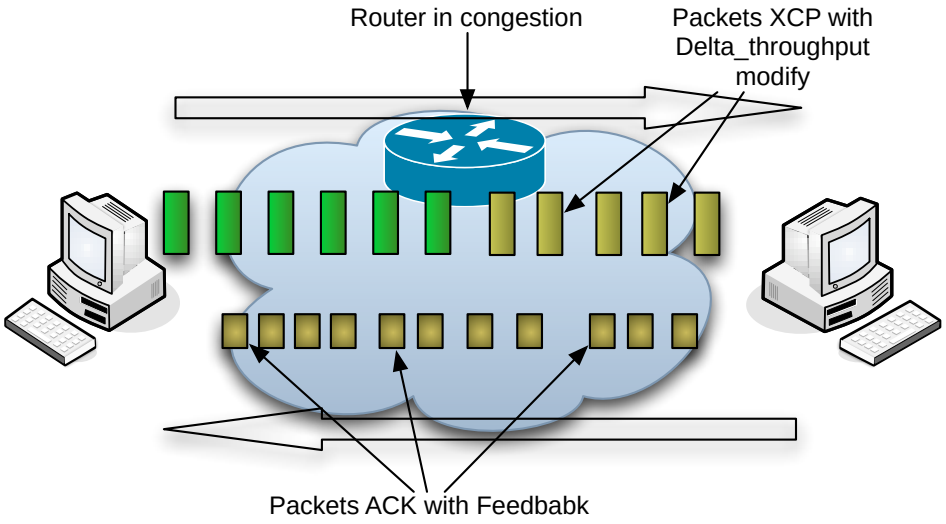

FIGURE 10: Principe de l'approche XCP

Cette approche, proposée dans [20], est une généralisation de l'approche ECN. En effet, selon l'approche ECN, un feedback est rendu à l'émetteur par l'intermédiaire d'une information binaire : la présence ou non d'une congestion.

Dans le cadre du XCP, l'échange d'information est plus important. Les paquets émis informent sur l'état de la connexion (débit, RTT). Les routeurs de coeur doivent agréger ces informations pour décider d'augmenter ou pas pas les débits de ces flux. Pour ce faire, la même méthode que l'ECN est adoptée, à savoir la modification d'un champ

du flux émis, retransmis par un champ de feedback sur les messages d'acquittements.

L'approche XCP apporte un certain nombre d'avantages face au protocole ECN :

- la réalisation d'une équité plus rapide.
- la possibilité de fournir des qualités différentes en jouant sur la valeur du Delta_throughput renvoyée dans le feedback.

2.2.5 *Evaluation dans le cadre des réseaux de transport mutualisé*

Nous avons précédement étudié plusieurs approches réalisant une gestion à l'échelle "microscopique" de la congestion. Nous avons mis en évidence un certain nombre de points communs qui ressortent de l'ensemble de ces travaux :

- tous ces travaux sont basés sur l'étude du réseau Internet. L'aspect mutualisation n'est traité uniquement que dans le cadre de nouveaux services à l'internaute (IpTV, VoIP), mais jamais abordé dans le cadre d'un réel réseau de transport mutualisé.
- ces approches déclarent la bande passante partagée "équitablement" quand chaque session protocoloraire utilise la même fraction du débit. Cette notion est fortement critiquable car elle ne prend pas en compte les aspects économiques du réseau. En effet, une session TCP n'a aucun rapport avec les entités économiques utilisant le réseau [10].
- les contraintes industrielles ne sont que partiellement prises en compte par ces solutions. L'association entre les mécanismes de coeur de réseau et les algorithmes d'extrémités ne répond pas aux contraintes sécuritaires. D'autre part, la politique de gestion des congestions occulte en partie la problématique du dimensionnement.

Ainsi une approche de type RIO peut être mise en oeuvre dans une architecture de transport mutualisé, mais sa gestion partielle de l'équité par session la rend d'une utilisation hasardeuse.

Quant à elles, les approches ECN et XCP ne correspondent pas aux contraintes d'un réseau de transport mutualisé. La régulation étant effectuée par un mécanisme distribué sur chaque émetteur, l'opérateur est contraint de faire confiance aux implémentations se trouvant chez ses clients. Ce qui n'est pas sans conséquence sur le plan sécuritaire.

Dans [30], P. Mosebekk met en évidence les difficultés de mise en oeuvre du protocole XCP dans un réseau. Dans [39], C. Wilson prouve

les possibilités de détournement du protocole pouvant entraîner des problèmes d'équité importants.

De plus la modification d'un champs de TCP est en complète contradiction avec le fonctionnement de MPLS qui ignore la couche 4 des paquets.

Finalement, l'étude de ces méthodes de gestion de la bande passante nous a permis de constater l'impossibilité d'utiliser simplement ces approches "microscopiques" pour réaliser une gestion plus fine de la bande passante résiduelle du coeur. En effet, deux difficultés apparaissent, d'une part la mise en oeuvre de ces approches réclame une modification de la chaîne complète du transport de l'émetteur au récepteur en passant par les noeuds intermédiaire, d'autre part les flux en mode non-connécté (type UDP) sont totalement ignorés par le mécanisme de partage. En conclusion, seul un système reliant les fonctionnalités de la planification réseau, à la gestion des phases de congestions devrait permettre de répondre à cette problématique.

Le chapitre suivant s'attachera à étudier, à l'aide de ces enseignements, un nouveau système de régulation dédié à cette double gestion.

3

PROPOSITION DE NOUVELLES APPROCHES

3.1 CONSIDÉRATIONS SUR LE NOUVEAU SYSTÈME

Nous avons vu précédemment la nécessité dans les réseaux de transport mutualisé de faire coexister et coopérer les approches "macroscopiques" de planification avec les approches "microscopiques" de gestion de la bande passante. Nous allons identifier dans un premier temps les éléments caractéristiques des réseaux de transport mutualisé qui n'ont pas été pris en considération dans les approches historiques.

3.1.1 *La base du réseau de transport mutualisé : le service*

Dans une architecture de réseau de transport mutualisé, le service est positionné sur une virtualisation de l'infrastructure. Chaque service d'un client repose sur un réseau virtuel. Cette séparation est uniquement protocolaire, la bande passante globale de l'infrastructure reste commune à tous. Un service se résume finalement à un PPVPN lui-même composé d'une part d'un certain nombre de tunnels MPLS qui sont à la base de la connectivité du service et d'autre part de noeuds virtuels qui émulent le comportement d'un équipement réseau réel.

A chaque service, on associe également des politiques de gestion des accès qui effectueront les opérations de policing, shaping et marking sur les flux clients.

3.1.2 *Une approche différenciant les tunnels*

La principale difficulté des précédentes approches étudiées au chapitre 2.2 réside dans leurs implémentations au niveau de la couche 4 du modèle OSI. En effet, la couche transport gère une connexion de bout en bout et ne permet pas de réaliser une optimisation locale à un secteur du réseau.

Ces observations nous amènent à la conclusion qu'il est nécessaire d'élargir les prérogatives des tunnels MPLS pour intégrer la notion de gestion de la bande passante au niveau du tunnel. Cela implique la possibilité d'imposer les bandes passantes des réseaux virtuels. La

réservation de bande passante effectuée par RSVP pour ces tunnels est un premier pas vers cette vision. Cependant, ce mode d'allocation de ressource ne permet pas une gestion dynamique de la bande passante.

Il convient alors d'intégrer dans notre système la faculté de mutualisation de la ressource. Un tunnel peut être traité face au réseau de transport mutualisé de la même manière qu'une connexion TCP est traitée face au réseau internet.

On retrouve d'ailleurs des propriétés similaires entre une connexion TCP et le débit présent dans un tunnel. En effet, le flux d'un tunnel n'est qu'un agrégat de connexions. Les propriétés comme l'élasticité sont alors transmises au flux agrégé. Dans le système proposé, nous utiliserons cette propriété à l'image des méthodes utilisées dans les approches microsopiques.

3.2 FORMALISATION DU SYSTÈME TDCN-ECN

Au regard des observations précédentes, nous proposons dans ce chapitre une nouvelle approche. Pour ce faire nous considérerons le tunnel comme l'élément de base du réseau.

Les études précédentes sur les approches de niveau 4 nous ont permis d'étudier un certain nombre de techniques probantes pour répondre à la problématique du partage équitable. Dans notre système, nous proposerons d'utiliser un retour d'information binaire à l'image du principe du protocole ECN. Lors de l'analyse des besoins de notre système, nous avons identifié un certain nombre de conditions à prendre en considérations :

- **la répartition du partage de la bande passante d'une liaison sera réalisée au prorata d'un contrat établi pour chaque tunnel,**
- **la signalisation permettant ce partage restera bornée au domaine de confiance de l'opérateur pour faciliter la phase d'industrialisation,**
- **l'objectif consistera à ne pas modifier les couches sessions afin de ne pas imposer aux clients le choix de ses protocoles.**

3.2.1 *Détection de la congestion*

Afin de réaliser une boucle de régulation efficace, il convient de détecter les périodes où une limitation est souhaitée. Pour ce faire, le système ECN effectue une détection de la congestion par le biais d'un marquage des paquets sur une bufferisation. Notre système étant

destiné à intervenir sur une congestion entre divers tunnels, il est essentiel que les flux associés aux tunnels reposent sur un même et unique buffer avant émission. La proposition suivante devra donc être respectée :

Règle TDCN 1 *Au sein des routeurs intermédiaires, l'ensemble des paquets des tunnels gérés par TDCN doivent transiter par un buffer unique avant la transmission sur le réseau.*

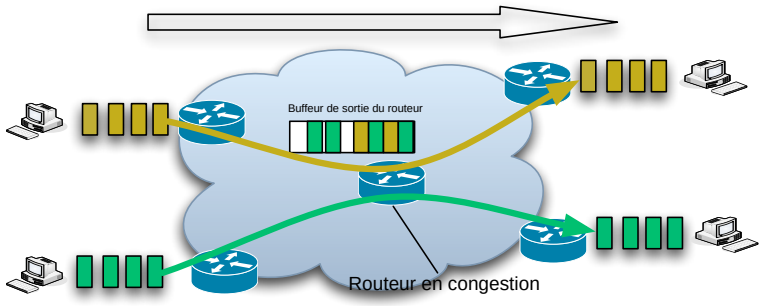

FIGURE 11: Mise en commun des buffeurs de sortie des routeurs de coeur

Un réseau best-effort peut répondre à cette proposition. De la même manière, un réseau mettant en oeuvre Diffserv [11] permet également l'utilisation du TDCN si on se satisfait de son usage sur une seule classe de service. Ceci autorisera une intégration plus aisée au sein des réseaux existants.

Le nouveau système se basera sur un remplissage partiel de ce buffer pour repérer une congestion. En effet, le principal indicateur d'une congestion vient des applications élastiques qui en essayant de croître entraînent une bufferisation sur le lien congestionné [4].

3.2.2 *Transport de l'information de la congestion sur les tunnels*

La technique retenue repose sur une limitation en accès lors des périodes de congestions. Ces congestions pouvant intervenir dans le coeur de réseau, il est nécessaire de rapatrier l'information de congestion sur le chemin du tunnel MPLS. Ces éléments nous amènent à établir la proposition suivante :

Règle TDCN 2 *Tout tunnel doit disposer d'un mécanisme permettant de connaître l'état de congestion du dit tunnel. Le tunnel est considéré congestionné si au moins une liaison du chemin du tunnel est congestionnée.*

La réutilisation des protocoles de signalisation des labels MPLS semble toute indiquée pour mettre en oeuvre ce mécanisme. En effet, le protocole RSVP-TE par exemple offre la possibilité de rajouter des paramètres [23]. Le maintien de sessions permet de conserver un rafraîchissement de l'état des paramètres pouvant descendre jusqu'à 300 ms.

3.2.3 *Fonctionnement des régulateurs d'accès "Dynamics Access Regulator" (DAR)*

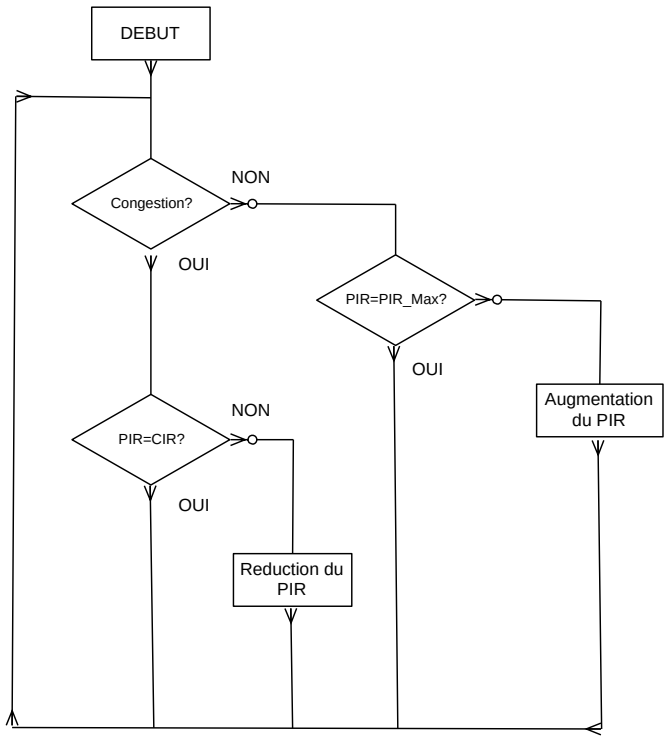

FIGURE 12: Algorithme du régulateur

A partir des deux règles définies précédemment, le système est capable de connaître sur le routeur d'accès du réseau l'état des tunnels

en aval. Grâce à cette information, chaque tunnel sera capable de réguler son débit en fonction de l'état du réseau. Pour ce faire, nous disposerons alors en amont de chaque tunnel d'un limitateur de débit que nous appellerons DAR. Il convient donc de définir la proposition suivante correspondant au fonctionnement du régulateur :

Règle TDCN 3 *Tout tunnel TDCN doit posséder en entrée un régulateur dédié prenant en compte l'information de congestion fournie par un protocole (règle 2). Ce régulateur dispose d'un paramètre définissant la limite maximale d'autorisation d'entrée dans le réseau (High Level Autorisation :HLA) et d'un paramètre définissant la limite minimale (Low Level Autorisation : LLA). La limite courante à un instant donné* k *est appelée* $CLA(k)$ *(Current Level Autorization).*

En complément de la règle précédente, il faut pouvoir garantir la convergence du régulateur en période de congestion. Ce qui signifie qu' il doit durcir sa politique d'autorisation d'entrée dans le réseau afin de ne pas participer à la congestion du coeur, mais au contraire de l'atténuer pour finalement la décaler à l'entrée du réseau. Ce fonctionnement peut être résumé dans la proposition suivante :

Règle TDCN 4 *Lors de la détection d'une congestion d'un tunnel, le régulateur doit diminuer sa CLA tout en respectant son minimum contractuelle. Par conséquent si la congestion dure suffisamment longtemps le CLA doit être égale au LLA. De plus le CLA ne descendra jamais son la valeur LLA.*

De la même façon, si aucune détection de congestion n'est effectuée sur le trajet d'un tunnel, la limitation doit être assouplie afin de permettre un maximum d'utilisation de la bande passante.

Règle TDCN 5 *Lors de la détection d'une non-congestion d'un tunnel, le régulateur doit augmenter sa CLA tout en respectant son maximum contractuelle. Par conséquent si le réseau ne congestionne pas pendant une période suffisament longue le CLA doit étre égale au HLA.*

3.2.4 *Respect du critère d'équité*

Le fonctionnement décrit précédemment revient à établir une nouvelle régulation qui vient alors s'ajouter à la régulation standard effectuée par les couches supérieures (TCP).La régulation effectuée au niveau de chaque tunnel aura un comportement dépendant uniquement des configurations du tunnel et non de la nature des flux le composant.

La mise en place de ces régulateurs va alors utiliser la propriété d'élasticité du flux agrégé mais restera indépendante de l'agressivité de ces flux. Il sera alors impossible pour un client d'influer sur le partage en modifiant par exemple son nombre de connexions TCP.

On pourra alors chercher à établir une équité indépendante de la nature des flux mais relevant de la nature des contrats commerciaux établis avec tel ou tel client.

La méthode pressentie pour obtenir cette équité contractuelle consiste à agir de manière proportionnellement aux contrats. Ainsi lors d'une période de congestion, un DAR disposant d'un haut contrat LLA, modifiera son CLA dans une moindre mesure qu'un régulateur disposant d'un plus faible contrat LLA. Inversement lors d'une période de non-congestion le tunnel disposant du meilleur contrat augmentera plus vite sa limite CLA.

On peut donc imaginer un critère d'équité dépendant par exemple du rapport LLA/CLA et ainsi effectuer la proposition suivante :

Proposition TDCN 6 *Afin de respecter le critère d'équité, un régulateur doit respecter :*

$$CLA(k+1) - CLA(k) = -f\left(\frac{LLA}{CLA(k)}\right)$$

si congestion

(3.1)

et :

$$CLA(k+1) - CLA(k) = g\left(\frac{CLA(k)}{LLA}\right)$$

si non congestion

(3.2)

avec :
f,g : des fonctions monotones croissantes positives.

D'autres propositions sont également envisageables et permettraient une convergence vers d'autres critères d'équité. Cependant, nous montrerons par la suite qu'un certain nombre de contraintes seront à respecter notamment sur les fonctions dérivées afin de ne pas rendre le système instable.

3.2.5 *Critère de stabilité du système*

Nous avons vu précédemment que le système va réduire ses limites d'accès lors d'une congestion jusqu'à la suppression de celle-ci. Une fois la congestion terminée, le système va pouvoir de nouveau élargir ses limites. Si la caractéristique des flux n'a pas évolué entre temps, la congestion réapparaîtra. Cela signifie que, dans le cas de profils de flux entraînant une période de congestion longue, le système oscillera entre congestion et non congestion au niveau du coeur de réseau. Il convient donc pour que l'amplitude d'oscillation du système soit faible de respecter la règle :

Règle TDCN 7 *La loi d'évolution doit diminué son CLA d'une valeur epsilonesque si la congestion perdure suffisamment longtemps*

3.2.6 *Détection de bande libre optimisée*

L'un des possibles effets sous optimal du système repose sur son temps de réaction qui peut être relativement lent notamment lorsqu'un profil de débit anciennement agressif devient passif. L'exemple classique est un transfert à fort débit finissant pour un client qui possède un fort LLA. Le transfert une fois terminé libère une part non négligeable de la bande passante de la liaison ; les autres clients ne peuvent alors l'utiliser du fait des limitations d'autorisations présentes sur leurs accès.

Afin d'atténuer le phénomène, il convient d'accélérer cette libération. Une longue période de non congestion devra se traduire par une suppression des anciennes limites en accès. Pour être efficace, le système devra accélérer la libération du CLA si aucune congestion n'apparait.

3.2.7 *Résumé du système*

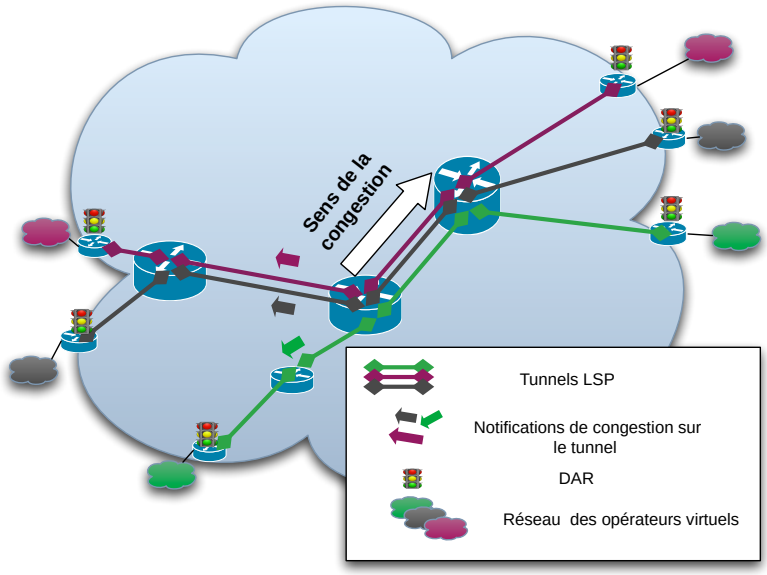

FIGURE 13: Principe du TDCN

Finalement, le système proposé repose sur l'association d'une signalisation de coeur et d'une régulation d'accès (figure 13). La complexité du système est principalement située dans les mécanismes du DAR et dans leurs interactions. Jusqu'à présent un certain nombre de critères ont été retenus en première approche mais non-démontrés. Dans le chapitre suivant nous nous focaliserons sur le fonctionnement du système et nous démontrerons l'intérêt de celui-ci.

4

ETUDE ANALYTIQUE DES PERFORMANCES DU SYSTÈME TDCN

4.1 MODÉLISATION DU SYSTÈME TDCN

Le chapitre précédent nous a permis de définir une proposition de mécanisme de régulation. Nous allons maintenant nous attacher à réaliser une modélisation de ce système. Dans un premier temps nous détaillerons le fonctionnement d'un régulateur et nous proposerons une modélisation. Puis nous focaliserons notre attention sur les interactions entre régulateurs.

4.1.1 *Le régulateur fractionnel*

Ce régulateur repose sur le principe d'actions intégrale. L'action du régulateur est dépendante des deux paramètres CLA et LLA. Par rapport à TCP, la variation TDCN est à une échelle temporelle supérieure (50 ms pour TCP, 1sec pour TDCN) et ainsi intervient en complément de celui-ci. Son action, dépendante de la valeur des contrats, va progressivement permettre d'établir une équité qui ne pourrait exister avec TCP seul. Lors d'une phase de congestion un régulateur modulera son action proportionnellement à sa CLA courante, mais également de manière inversement proportionnelle à sa LLA contractualisée pour avantager ainsi les contrats les plus forts (Eq. 3.1).

Lors d'une phase de non congestion le raisonnement inverse est appliqué(Eq. 3.2).

Dans la pratique, le décrément en congestion est proportionnel au rapport $\frac{CLA}{LLA}$ et l'incrément en non-congestion est proportionnel au rapport $\frac{LLA}{CLA}$.

Evolution du régulateur en congestion

En congestion, la valeur à l'instant k notée $D(k)$ du décrément sera calculée grâce à la formule $D(k) = A * \frac{CLA(k)}{LLA}$ où A est une constante exprimée en bits par secondes.

Pour le moment la valeur de la constante A est attribuée arbitrairement. Nous verrons par la suite que la valeur de cette constante peut

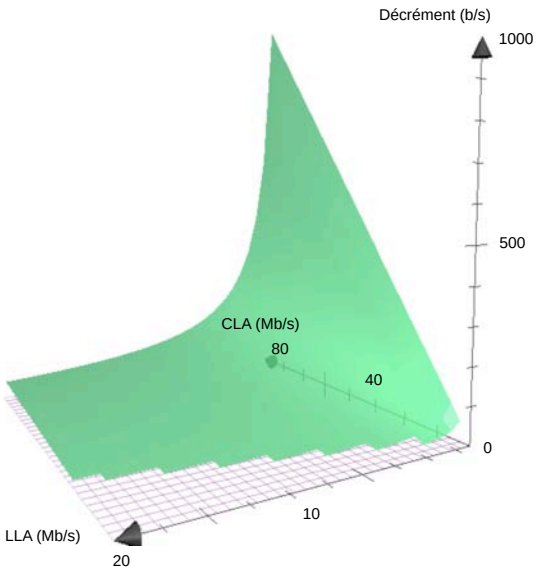

FIGURE 14: Valeur du décrement (A :15b/s)

être optimisée. La figure 14 représente l'évolution de la valeur de ce décrément en fonction des CLA et LLA. On observe ainsi dans le diagramme en 3D qu'à CLA constant le plus faible contrat décrémentera plus vite et que pour un même contrat le régulateur à la limite la plus forte décrémentera plus rapidement.

Durant une période de congestion, on pourra déterminer l'évolution d'un régulateur grâce à la règle récursive suivante, comme le montre la figure 15 :

$$CLA(k+1) = CLA(k) - \frac{CLA(k)}{LLA} * A \tag{4.1}$$

Evolution du régulateur en non-congestion

En non-congestion, la valeur à l'instant k notée $I(k)$ de l'incrément sera calculée avec la formule $I(k) = B * \frac{LLA}{CLA(k)}$ où B est une constante exprimée en bits par seconde. Là encore, nous étudierons plus tard l'optimisation de cette constante. La Figure 16 représente la valeur de cette incrément en fonction du CLA et du LLA.

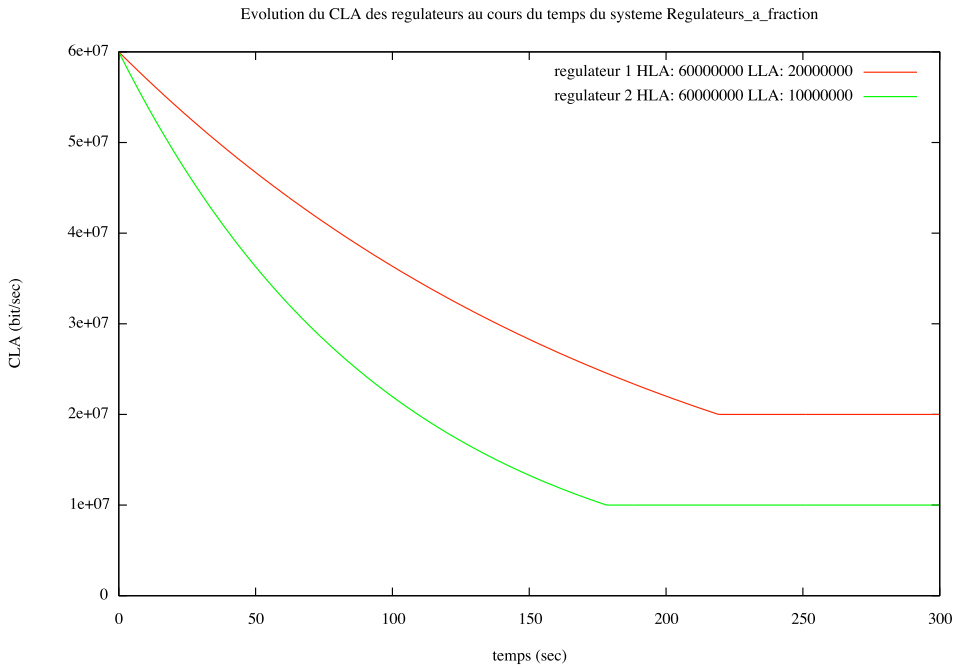

FIGURE 15: Evolution au cours du temps de la CLA en congestion(A=100)

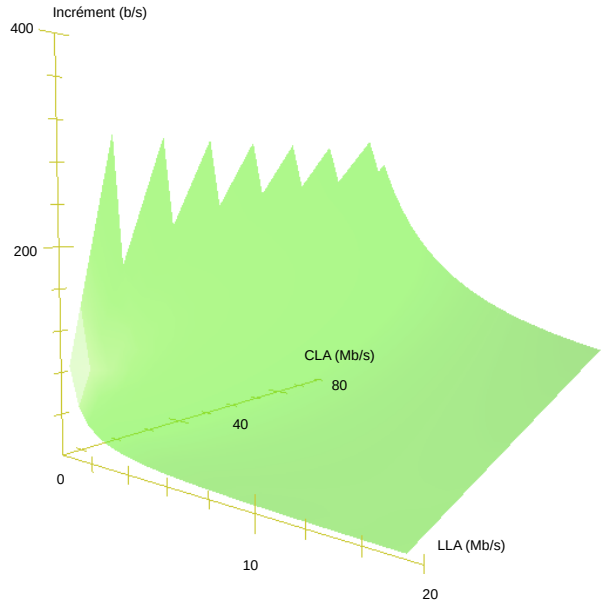

FIGURE 16: Valeur de l'incrément (A :15b/s)

Durant cette période, on pourra également déterminer l'évolution du régulateur grâce à la formule suivante, comme le montre la figure 17 :

$$CLA(k+1) = CLA(k) + \frac{LLA}{CLA(k)} * B \tag{4.2}$$

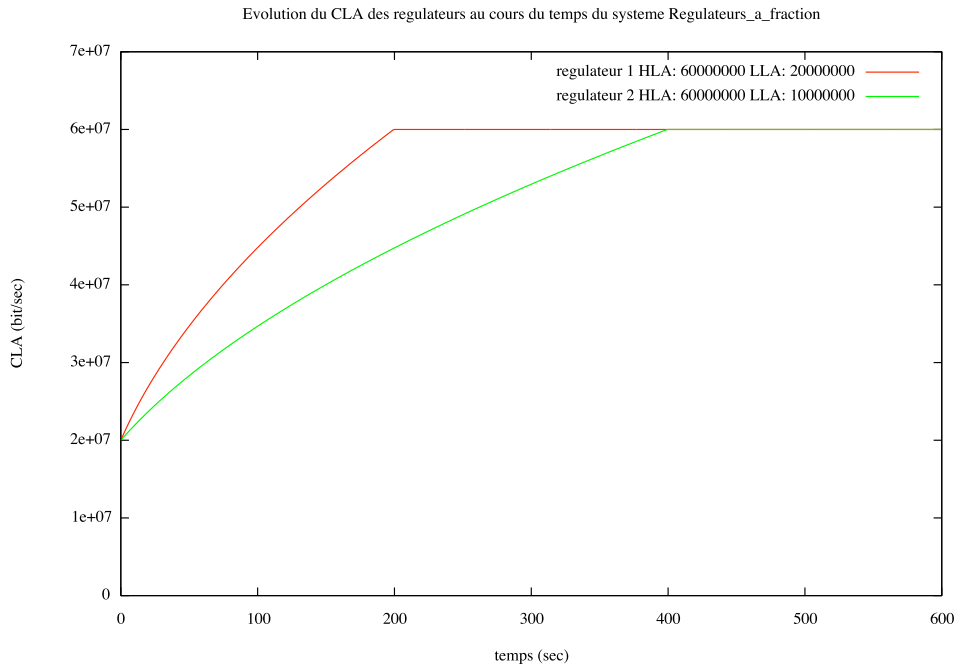

FIGURE 17: Evolution au cours du temps du CLA en non-congestion (B=400)

4.1.2 *Mise en équations*

Modélisation d'un DAR

Lors d'une période de congestion un régulateur agit comme une suite. En effet, son fonctionnement récursif justifie ce type de modélisation. Le régulateur répond à l'équation 4.3 en congestion et à l'équation 4.4 en non-congestion.

$$CLA(k+1) = max(CLA(k) - f(CLA(k)), LLA) \tag{4.3}$$

$$CLA(k+1) = min\left(CLA(k) + g\left(CLA(k)\right), HLA\right) \tag{4.4}$$

En temps normal plusieurs régulateurs (n), possédant une limite CLA, noté CLA_i pour le ième régulateur, traverse le lien congestionné. En non-congestion, la somme de leur limite peut dépasser la capacité des liaisons.

Lorsqu'une liaison est congestionnée, la somme des débits clients sera égale à la capacité de la liaison. Le système TDCN diminuera l'ensemble des CLAs, jusqu'à l'obtention d'une limitation des débits. On obtient alors une valeur E égale à la somme des CLA.

Dans ces conditions, la congestion de coeur va perdurer tant que la somme des CLA des tunnels est supérieure à la valeur E. Dès qu'elle passera en dessous, les flux ne pourront plus maintenir la congestion.

On remarquera toutefois que la valeur E peut être égale à la bande passante dans le cas particulier où tous les clients utilisent des flux élastiques susceptibles de prendre toute la bande passante.

Par conséquent, pour une liaison de Débit E composée de n DARs, on aura une congestion lorsque l'équation 4.5 est respectée et une non-congestion lorsque l'équation 4.6 est respectée.

$$\sum_{k=0}^{n} CLA(k) > E\ ,\ \forall k \tag{4.5}$$

$$\sum_{k=0}^{n} CLA(k) \leqslant E\ ,\ \forall k \tag{4.6}$$

L'ensemble des observations précédentes repose sur un certain nombre d'hypothèses :

- le temps de transport de l'information de congestion est négligeable face au délai entre deux décréments,
- les régulateurs agissent tous au même rythme et l'effet de leurs désynchronisations est négligeable,
- les flux élastiques agrégés occupent instantanément la bande passante disponible.

On peut ainsi poser le système suivant dérivé des conclusions de la section :

$$\begin{cases} CLA_i(k+1) = & max\,(CLA_i(k) - f(CLA_i(k), LLA_i, HLA_i)) \\ & \forall i \text{ et si : } \sum_{i=0}^{n} CLA_i^k > E \\ CLA_i(k+1) = & min\,(CLA_i(k) + g(CLA_i(k), LLA_i, HLA_i)) \\ & \forall i \text{ et si : } \sum_{i=0}^{m} CLA_i(k) \leqslant E \end{cases} \tag{4.7}$$

4.1.3 *Etude d'un système à DAR fractionnel*

A partir du modèle proposé précédemment, nous allons effectuer une simulation mettant en évidence les différentes phases d'une congestion TDCN.

Nous considérons une liaison congestionnée regroupant 3 tunnels MPLS managés avec notre nouveau système TDCN. La liaison possède une capacité de 1 Gb/s. Les configurations TDCN sont identiques pour chacun des tunnels (HLA=1 Gb/s LLA= 100 Mb/s). Grâce à la modélisation que nous venons de réaliser, nous pouvons prédire le fonctionnement du système lorsque l'ensemble des clients va essayer de consommer les 1 Gb/s de données de la HLA.

La figure 18 est un exemple de l'évolution de CLAs de trois clients mis en concurrence suite à l'apparition de cette congestion. Chaque système dispose de courbes propres en fonction d'une part du paramètrage des régulateurs mais également en fonction des **conditions initiales**. Nous étudierons leurs impacts respectifs dans un chapitre ultérieure. Cependant, d'une manière générale, on retrouve trois phases distinctes suite à l'apparition d'une congestion dans un système TDCN :

- **La zone de congestion :** Cette zone apparaît au début de la congestion de coeur. Pendant cette période, tous les régulateurs durcissent pas à pas leurs politiques d'accès afin de supprimer cette congestion. Une fois la congestion stoppée (passage de la somme des CLA_i au dessous de la valeur E), le système rentre dans son deuxième état.
- **La zone d'équilibrage :** Pendant cette période, le système oscille entre un état de congestion et un état de non congestion (oscillation autour de la somme des CLA_i autour de E). A l'issue de la phase précédente, la CLA des régulateurs s'est positionnée dans un état intermédiaire dépendant des configurations mais également des conditions initiales. Par la suite les régulateurs

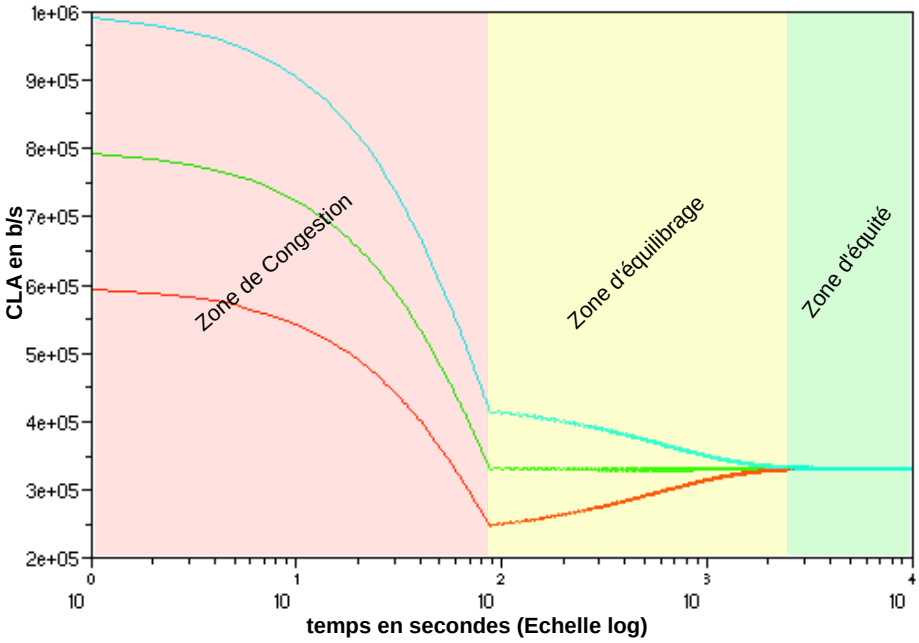

FIGURE 18: Système simple à base de DAR fractionnels

vont modifier leur CLA en faisant varier leurs incréments et décréments pour finalement atteindre un état stable indépendant des conditions initiales.

– **La zone d'équité :** Il s'agit du point de fonctionnement stable du système. Nous montrerons par la suite que cet état correspond à l'équité de partage de la bande passante entre les différents tunnels. Cet état se maintiendra tant que les flux élastiques à l'intérieur des tunnels réclameront de la bande passante. Au niveau du lien précédemment congestionné, celui-ci oscille maintenant entre un état de congestion de coeur et de non congestion.

Etude de la période de congestion

La durée de cette période dépend de deux facteurs : le premier est la "distance" entre les conditions initiales et le point d'équilibre et le second est le paramètre A du DAR.

D'autre part, plus la somme des HLA des régulateurs dépassera la bande passante des liaisons, plus les délais pour couper une congestion seront importants. Ainsi on définira un taux d'over-provisioning d'une liaison comme étant le rapport entre la somme des HLA des tunnels et la bande passante de la liaison.

Dans tout les cas, il est intéressant de noter qu'il est possible de prédire le temps maximal que mettra le système pour supprimer cette congestion. En effet, pendant cette période les performances réseaux sont altérées ; on observe un taux de drops important et une forte augmentation de la latence.

Pour diminuer ce délai, il faut augmenter la vitesse de décrémentation des DARs en augmentant le paramètre A (equa.4.1.1). Or nous verrons par la suite que l'augmentation de cette valeur peut entraîner des effets néfastes. Un compromis devra alors être trouvé.

Par contre, cette zone est caractérisée par une succession de décréments. On peut ainsi améliorer la vitesse de convergence en conservant en mémoire le nombre de décréments successifs et en appliquant une pénalité supplémentaire au décrément et ainsi passer plus rapidement à l'étape suivante d'équilibrage.

Etude de la période d'équilibrage

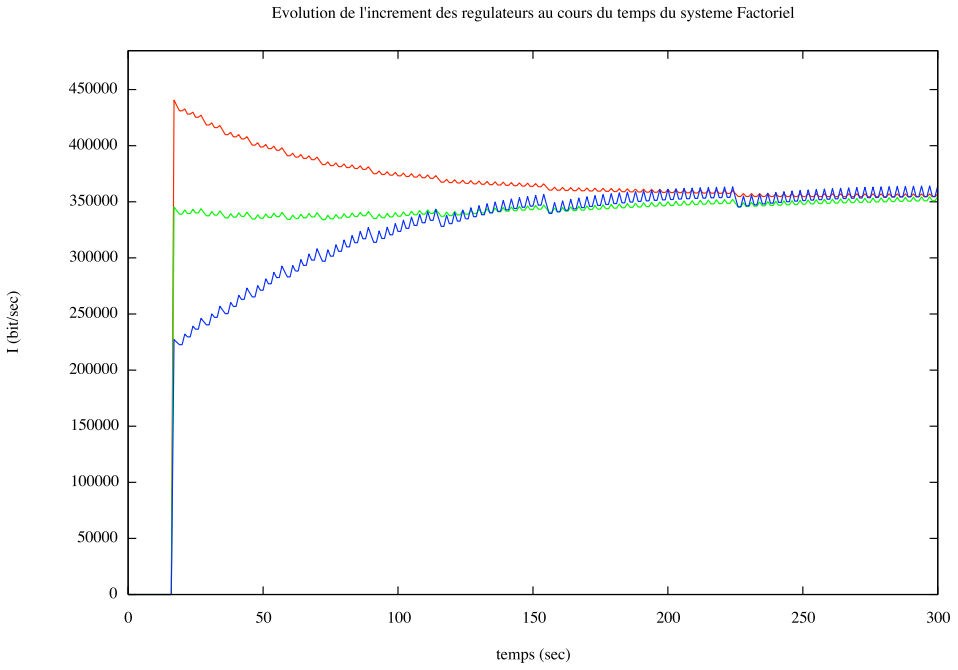

FIGURE 19: Evolution de l'incrément d'un système composé de 3 régulateurs

Pendant cette période, les régulateurs basculent alternativement des modes de congestion aux modes de non-congestion. Les différences entre les incréments et les décréments de chaque régulateur vont modifier la manière dont s'effectue le partage.

Lors de la phase précédente, les régulateurs vont diminuer leurs CLA jusqu'à une valeur dépendante des conditions initiales. La pondération étant fonction des CLA courantes, les décalages vont se modifier pour tendre vers une valeur de décalage unique et commune à tous les régulateurs.

La figure 19 montre l'évolution des incréments d'un système à trois régulateurs. Le résultat est l'obtention d'un point équilibre unique où chaque régulateur évoluera selon le même incrément et décrément.

A noter que la valeur de l'incrément et du décrément n'est pas forcement identique. Dans le cas d'une asymétrie, le système restera durant plusieurs pas consécutifs dans le même état avant de faire une bascule. Ce phénomène est très intéressant car il permet ainsi de conserver le système en état de non-congestion plus longtemps entraînant ainsi de meilleurs performances réseaux.

Etude de la période d'équité

La section précédente nous permet de constater que tous les DARs finissent par avoir un incrément moyen identique, ainsi qu'un décrément moyen identique. Ce résultat va nous permettre de prévoir le partage qui sera effectué entre les régulateurs.

En effet, on peut poser les équations suivantes pour le ième régulateur, nous noterons $\overline{CLA_i}$ la valeur $CLA_i(k)$ lorsque l'équité est atteinte :

$$\begin{array}{l|l} A * \frac{\overline{CLA_i}}{LLA_i} = \overline{D} & B * \frac{LLA_i}{CLA_i} = \bar{I} \\ \Rightarrow CLA_i = \frac{\overline{D} * LLA_i}{A} & \Rightarrow CLA_i = \frac{B * LLA_i}{\bar{I}} \end{array} \tag{4.8}$$

Or nous savons également que la somme des CLAs lors de de l'équité est égale à la bande passante de la liaison congestionnée (Bp) par conséquent on a :

$$\begin{array}{l|l} \sum_{i=0}^{n} \overline{CLA_i} = E & \\ \sum_{i=0}^{n} LLA_i * \frac{\overline{D}}{A} = E & \sum_{i=0}^{n} LLA_i * \frac{B}{\bar{I}} = E \\ \frac{\overline{D}}{A} = \frac{E}{\sum_{i=0}^{n} LLA_i} & \frac{B}{\bar{I}} = \frac{E}{\sum_{i=0}^{n} LLA_i} \end{array} \tag{4.9}$$

On a alors pour chaque régulateur le respect de l'équation suivante :

$$\overline{CLA_i} = E * \frac{LLA_i}{\sum_{i=0}^{n} LLA_i} \tag{4.10}$$

On peut en conclure que le partage en mode congestionné de la bande passante est un partage barycentrique où la LLA fait office de poids.

La modélisation réalisée dans ce chapitre, nous a permis de prédire le fonctionnement du système lors de l'apparition d'une congestion. Jusqu'à présent celle-ci nous a servi à identifier les trois phases symptomatiques du système.

Par la suite, nous utiliserons cette modélisation et les résultats obtenus pour établir et tester un système exploitable dans un cas réel.

4.1.4 *Etude d'une application basée sur le TDCN*

Le framework TDCN a de multiples façons d'être utilisé et configuré. Le choix retenu dépendra notamment du modèle commercial mais également des choix industriels et des paramètres que l'on souhaitera optimiser.

Dans ce chapitre, nous allons proposer une manière spécifique de configurer ce framework et nous montrerons toutes les garanties qui en découlent.

L'approche par métaux précieux

Afin de simplifier les possibilités, nous proposerons dans ce chapitre trois types de contrats standards proposés par les opérateurs : Gold, Silver et Bronze. Contrairement aux offres actuelles, nous verrons que les garanties fournies ne reposent pas sur des approches statistiques mais sont basées sur le fonctionnement intrinsèque du système.

Exemple illustratif

Afin d'illustrer ce cas pratique, nous allons nous focaliser sur une liaison 1 Gb/s dans laquelle on situe :

- 10 contrats gold à 10Mb/s,
- 60 contrats silver à 10Mb/s,
- 100 contrats bronze à 10Mb/s,
- 1 contrat gold à 100Mb/s,
- 1 contrat silver à 100Mb/s,
- 1 contrat gold à 100Mb/s.

Le débit des contrats se traduit directement dans le système par la valeur HLA des régulateurs. On remarquera dans le cas présent que la somme des HLA est supérieure à la capacité de la liaison. On a donc un taux d'under-provisioning de 2. Dans cette situation, on considère que nos clients n'utilisent pas simultanément l'ensemble de leur bande passante. Le contrat ne garantit de toute manière qu'une partie de cette bande passante.

La configuration du système TDCN

Configuration du LLA

Pour réaliser la différentiation entre les différents types de contrats, nous fournissons un LLA au prorata (noté α) du contrat HLA. Ainsi, les clients Gold auront 70% de leur trafic garanti, les Silver 50 % et les Bronze 20%.

Ainsi dans notre exemple illustratif, l'ensemble du débit garanti s'élève à 710Mb/s. Ceci correspond à une liaison disposant pour le trafic garanti un taux d'over-provisioning de 40%. On aura donc ici la certitude de pouvoir fournir les débits contractuels.

Dans cette approche, on démontre d'une part que même le contrat le plus bas dispose d'une garantie, et d'autre part que la valeur du contrat reste modulable en fonction des besoins de chaque client.

Configurations des paramètres A et B

Afin de choisir la méthode de configuration des paramètres A et B, nous étudions l'évolution des CLA lors de l'apparition d'une congestion dans notre exemple illustratif (section 4.1.4) avec des valeurs de A et B égales à 100 (fig. 20). La figure 20 met en évidence que chaque type de contrat dispose d'une bande passante au prorata de sa HLA et ce quelque soit la valeur de la HLA. Par contre, nous pouvons observer que l'équité est plus longue à obtenir pour les contrats de 100Mb/s. Afin de palier à cela, nous proposerons une valeur de A et B au prorata (noté β) du contrat HLA. Ainsi, une configuration de A et B au 1/100 de la HLA donne le fonctionnement suivant (fig. 21). Cette modification permet à l'ensemble des comportements des régulateurs d'agir proportionnellement à leurs propres HLA.

4.1.5 *Impact de l'étude temporelle*

Les travaux effectués sur un cas pratique, nous ont permis de mettre en évidence un certain nombre d'avantages du système. Cependant,

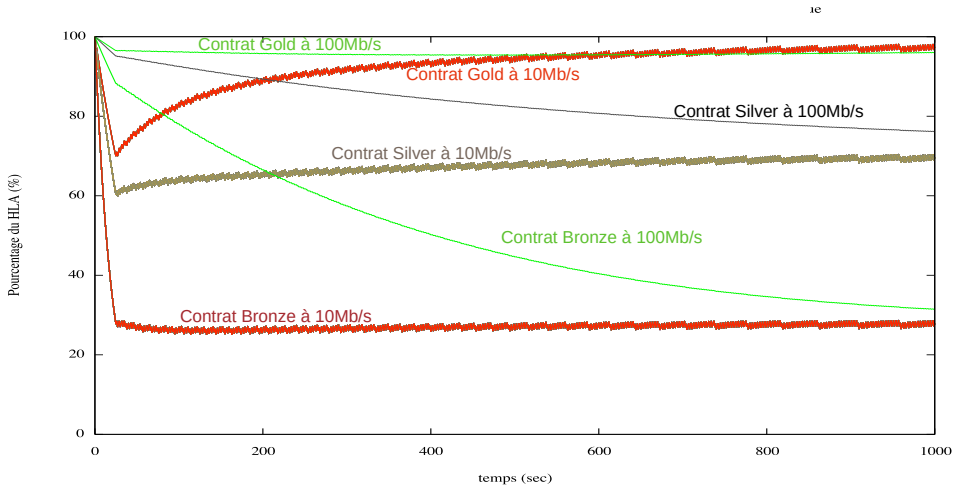

FIGURE 20: Evolution de la CLA d'un système à descente constante(en % de la HLA)

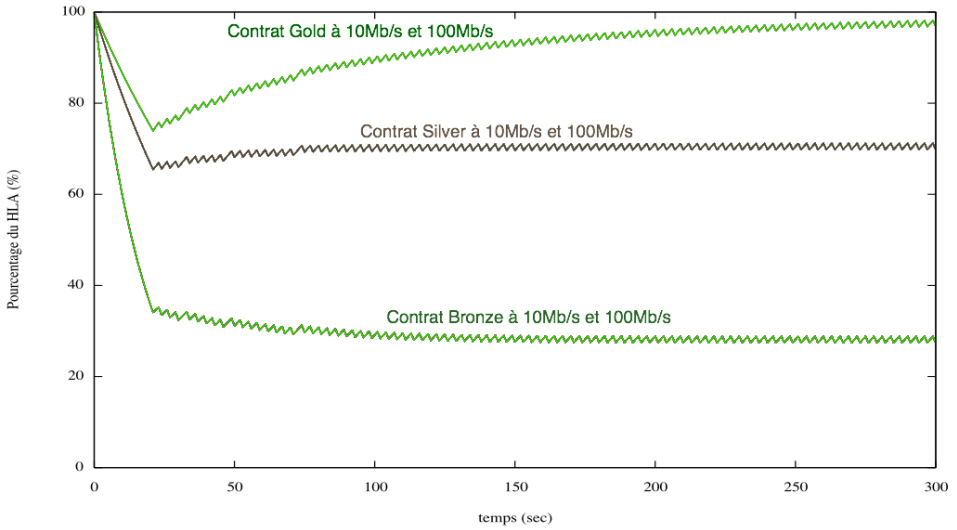

FIGURE 21: Evolution de la CLA d'un système à descente variable(en % de la HLA)

cette étude temporelle ne permet pas d'acter les choix optimaux des constantes ou de certifier son bon fonctionnement systémique.

Nous pouvons observer que le système fournit à tout moment un CLA supérieur pour chaque client au contrat minimum même lors de la phase de congestion.

Dans le chapitre suivant, nous proposerons une étude par diagramme de phase. Celle-ci nous permettra de proposer et valider des méthodes de choix pour les diverses constantes du système.

4.2 ETUDE PAR DIAGRAMME DE PHASE DE LIAISON TDCN EN CONGESTION

L'étude temporelle du système nous a permis précédemment d'identifier différentes phases de fonctionnement. Dans ce chapitre, nous allons approfondir l'étude de l'évolution de l'état du système en générant des indicateurs. Ceux-ci serviront par la suite de base de comparaison pour étudier l'impact de paramètres tels que l'établissement des constantes, les conditions initiales, le nombre de régulateurs, la variété des contrats,...

4.2.1 *Etude d'un système simple à deux régulateurs*

Définition du système

Dans un premier temps, nous considérons un système ne comprenant que deux régulateurs. Nous généraliserons ensuite la méthode pour des systèmes comprenant plus de régulateurs. L'étude que nous nous proposons de faire se base sur la création de diagramme de phase. Dans un premier temps nous définissons le vecteur d'état. Celui-ci sera logiquement le vecteur de dimension 2 composé des valeurs des deux CLA de chaque régulateur. On notera de façon générale le vecteur d'état C. Nous étendrons, par la suite, la notation vectorielle en notant le vecteur L le vecteur des LLAs et les vecteurs A et B seront les paramètres des régulateurs (nous considérons les régulateurs fractionnels décrits dans le chapitre 4.1.1). On a donc :

$$C(k)\begin{pmatrix} CLA_0(k) \\ CLA_1(k) \end{pmatrix}, \quad \vec{L}\begin{pmatrix} LLA_0 \\ LLA_1 \end{pmatrix}, \quad \vec{H}\begin{pmatrix} HLA_0 \\ HLA_1 \end{pmatrix}, \quad \vec{A}\begin{pmatrix} A_0 \\ A_1 \end{pmatrix}, \quad \vec{B}\begin{pmatrix} B_0 \\ B_1 \end{pmatrix} \tag{4.11}$$

On note $C_i(k)$ la iéme composante de $C(k)$. C'est à dire on a $C_i(k) = CLA_i(k)$.

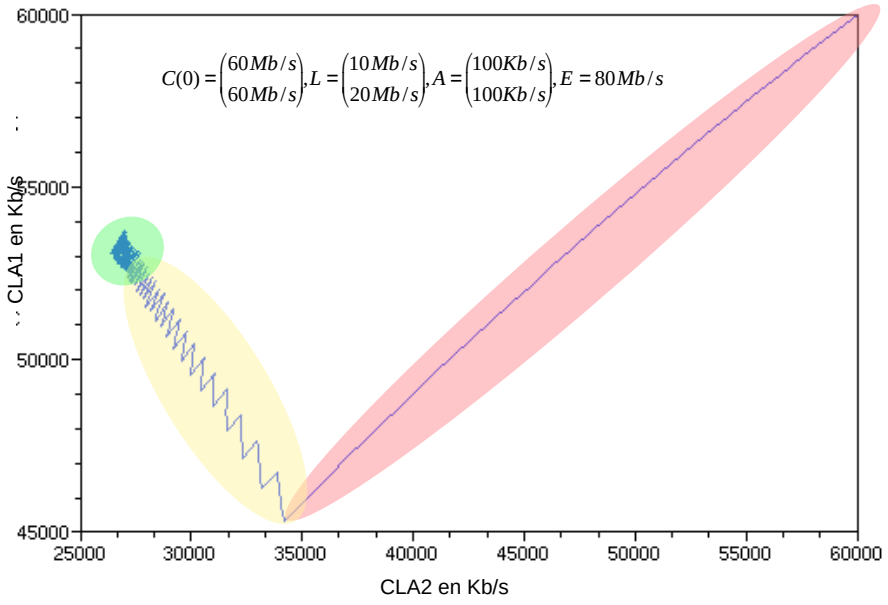

FIGURE 22: Diagrammes de phases d'un système à 2 régulateurs

Diagramme de phase

A l'aide de la modélisation réalisée dans le chapitre précédent, on peut réaliser le diagramme de phase du système. La figure 22 est un exemple de ces diagrammes.

L'évolution du système met en évidence les trois phases observées lors de l'étude temporelle à savoir :

- La première phase (appelée précédemment phase de congestion) caractérisée par une chute de l'ensemble des valeurs des CLAs,
- une seconde phase (dite d'équilibrage) où les valeurs CLA_i varient fortement mais où la somme des CLA_i oscille autour d'une valeur E (Valeur de coupure de la congestion cf. 4.1.2) ,
- une dernière phase (dite équilibrée) où l'ensemble des CLA converge autour d'un point fixe. Celui étant déterminé en fonction du vecteur L et de la valeur E (Démontré dans 4.1.3).

Test d'indépendance aux conditions initiales

La représentation en diagramme de phase nous a permis d'observer les trois phases de la convergence. Cependant, celui-ci a été réalisé pour

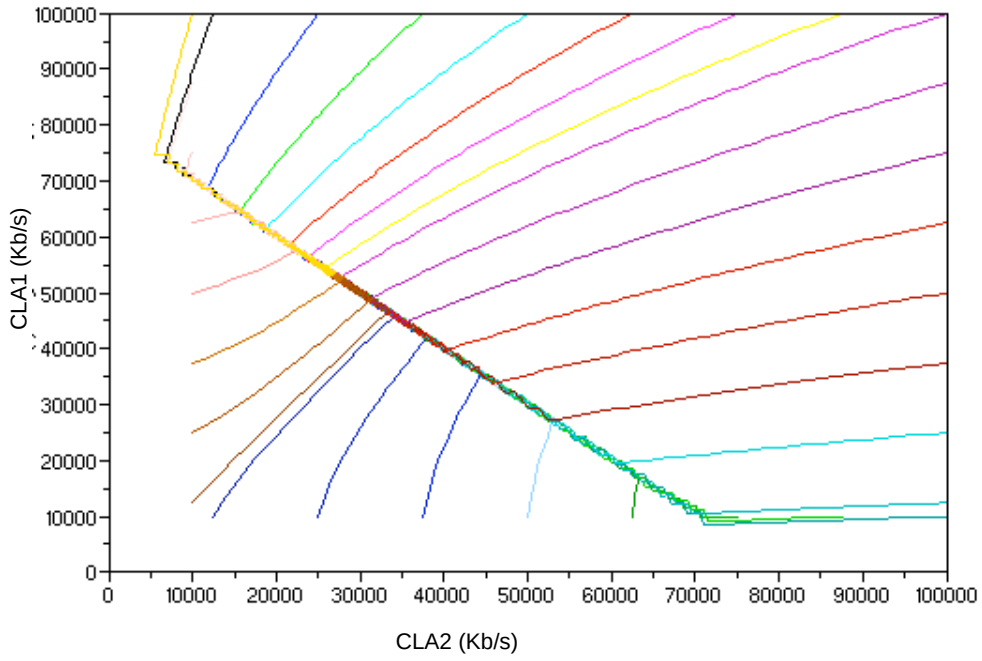

FIGURE 23: Diagrammes de phases d'un système à 2 régulateurs à différentes condition initial

une condition initiale particulière. On peut d'ores et déjà s'interroger sur l'impact de ces conditions.

Afin de contrôler le fonctionnement du système, nous réalisons le diagramme de phase de systèmes identiques mais possédant des conditions initiales sur le CLA différentes.

Le résultat de ce test (Figure 23) met en évidence d'une part que le point d'équilibre est invariant et d'autre part on remarque que la phase d'équilibrage de tous les systèmes se trouve être une oscillation autour d'une droite unique ayant pour équation 4.12 :

$$C_0(k) + C_1(k) = E$$

pour $k > k_0$ avec k_0 l'instant où la trajectoire atteint la phase d'équilibrage. (4.12)

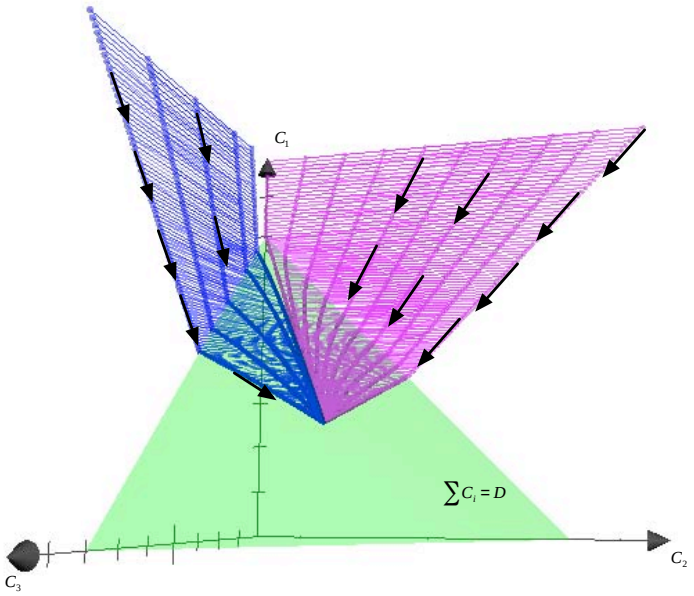

FIGURE 24: Diagramme de phase de systèmes à 3 régulateurs avec différentes conditions initiales

4.2.2 *Généralisation aux systèmes à N régulateurs*

Cas du système à 3 régulateurs

Afin d'étendre les observations effectuées dans le système à deux régulateurs, nous allons poursuivre l'étude par un système de dimension trois. Celui-ci est défini par les vecteurs de l'équation 4.13 :

$$C(k)\begin{pmatrix} CLA_0(k) \\ CLA_1(k) \\ CLA_2(k) \end{pmatrix}, \quad L\begin{pmatrix} LLA_0 \\ LLA_1 \\ LLA_2 \end{pmatrix}, \quad H\begin{pmatrix} HLA_0 \\ HLA_1 \\ HLA_2 \end{pmatrix}, \quad A\begin{pmatrix} A_0 \\ A_1 \\ A_2 \end{pmatrix}, \quad B\begin{pmatrix} B_0 \\ B_1 \\ B_2 \end{pmatrix} \tag{4.13}$$

Dans ce cas de figure (figure 24), on peut encore observer les trois phases. Par contre, la droite $\mathcal{D}$ devient un plan

De la même façon on peut vérifier une nouvelle fois la véracité du partage barycentrique à l'état d'équilibre.

Observation sur les systèmes de dimension supérieure

Etude de la Notion de "distance" L'observation des diagrammes de phase ne peut être effectuée sur les systèmes de degrés supérieurs. Cependant, si nous connaissons la valeur E, nous pouvons pré-calculer le point de convergence grâce à l'équation 4.10. A partir de là, nous pouvons définir une fonction de "distance" entre le point de fonctionnement observé au temps t et le point de convergence. Cette fonction est définie par l'équation 4.14.

$$Distance(k) = \sqrt{\sum_{i=1}^{n} (C_i(k) - C_i^{conv})^2} \tag{4.14}$$

$$Distance(k) = \sqrt{\sum_{i=1}^{n} (C_i(k) - E * \frac{LLA_i}{\sum_{i=1}^{n} LLA_i})^2} \tag{4.15}$$

L'intérêt de cette fonction repose sur la possibilité de réaliser un graphique d'évolution au cours du temps de tout système quelque soit sa dimension. Pour une représentation plus visuelle, nous effectuerons un affichage logarithmique plus propice à mettre en évidence les trois phases du système (figure 25).

traitement numérique et diagramme de synthèse

Afin de faciliter les traitements numériques et la compréhension du fonctionnement du système étudié, nous effectuerons un lissage de la courbe en réalisant une moyenne mobile. La courbe obtenue est alors représentative d'un système défini par les paramètres L,A et B face à l'événement apparition d'une congestion ayant pour débit d'équilibre E avec comme condition initiale le vecteur état instantané C_0. Cette courbe (figure 26) possède finalement deux caractéristiques essentielles pour notre système :

- **L'erreur de convergence** $distance_\infty$ **:** valeur moyenne de la distance lorsque k tend vers l'infini.
- **Délai de convergence** k_c **:** représenté par le temps que le système a mis pour arriver à 10% de son erreur de convergence minimale.

Par la suite, nous utiliserons ces indicateurs pour comparer l'efficacité sur des systèmes configurés différemment.

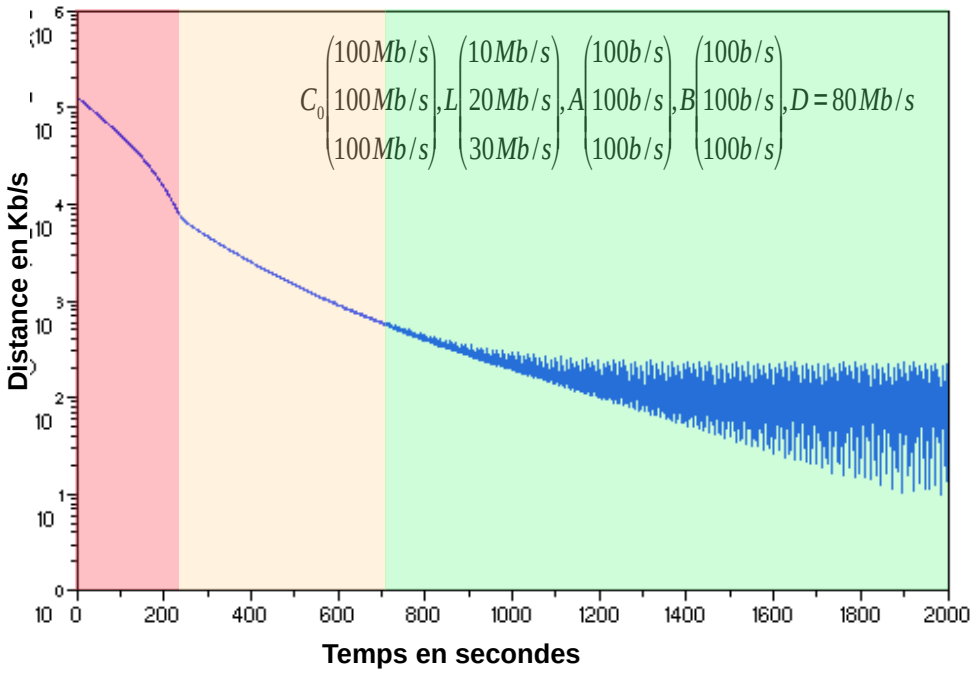

FIGURE 25: Evolution de la distance de convergence d'un système

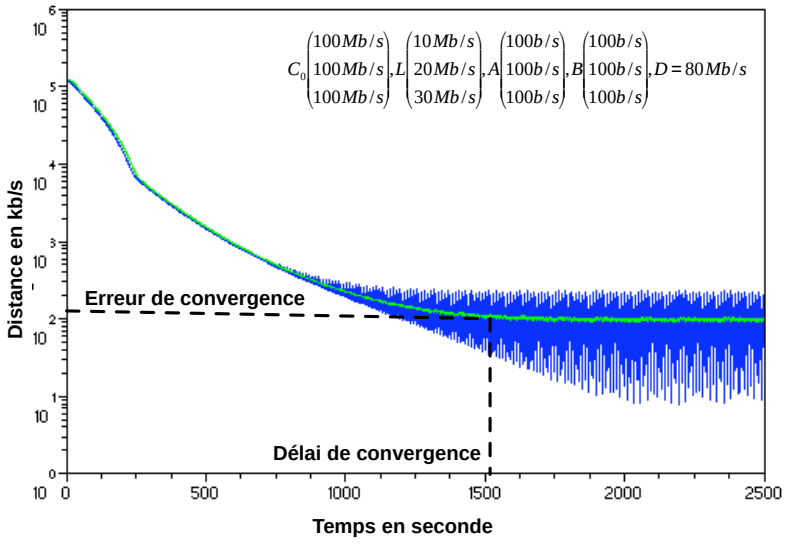

FIGURE 26: Evolution de la distance fenêtrée d'un système

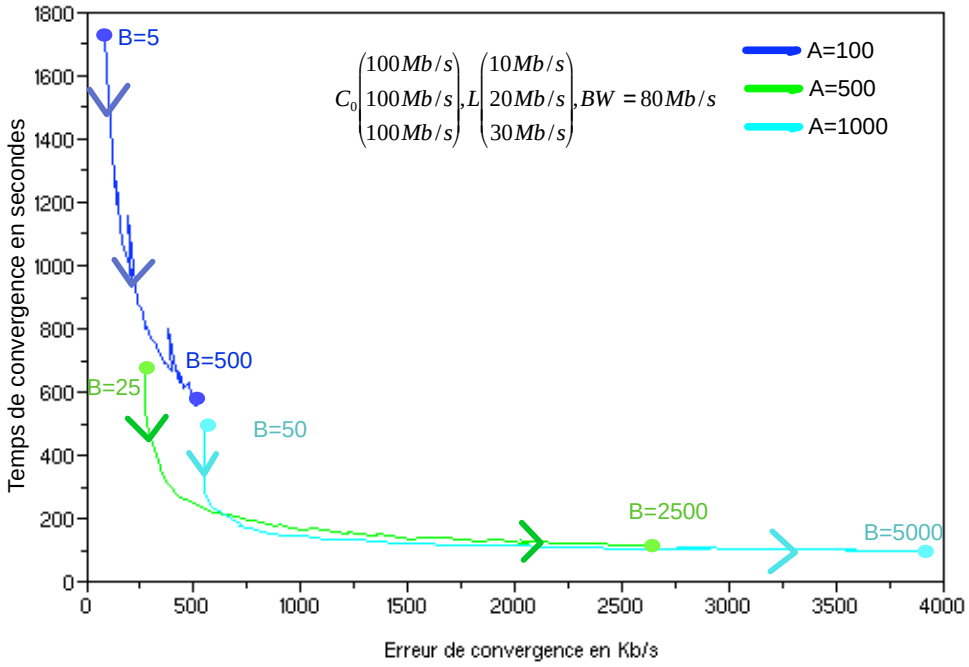

FIGURE 27: Evolution du système en fonction des paramètres A et B

4.2.3 *Analyse de liaison sur les Systèmes TDCN*

Etude des effets des paramètres A et B

Lors de l'établissement d'exemples, nous avons jusqu'à présent choisi les paramètres A et B de manière empirique. Grâce aux résultats du chapitre précédent, nous pouvons comparer l'efficacité sur des systèmes configurés différemment.

La figure 27 représente un système identique avec des valeurs A et B variantes. On peut constater que la précision de la convergence est directement corrélée avec ces valeurs. Ainsi, plus A et B possèdent des valeur élevés, plus le système converge rapidement par contre l'erreur de convergence va également croître. On a donc un compromis à trouver pour la valeur du couple de paramètres A et B afin d'obtenir une convergence satisfaisante dans un temps raisonnable.

Effet du nombre de régulateurs

Un des questionnements que peut susciter notre système est la possibilité que possède une liaison de supporter un nombre variable

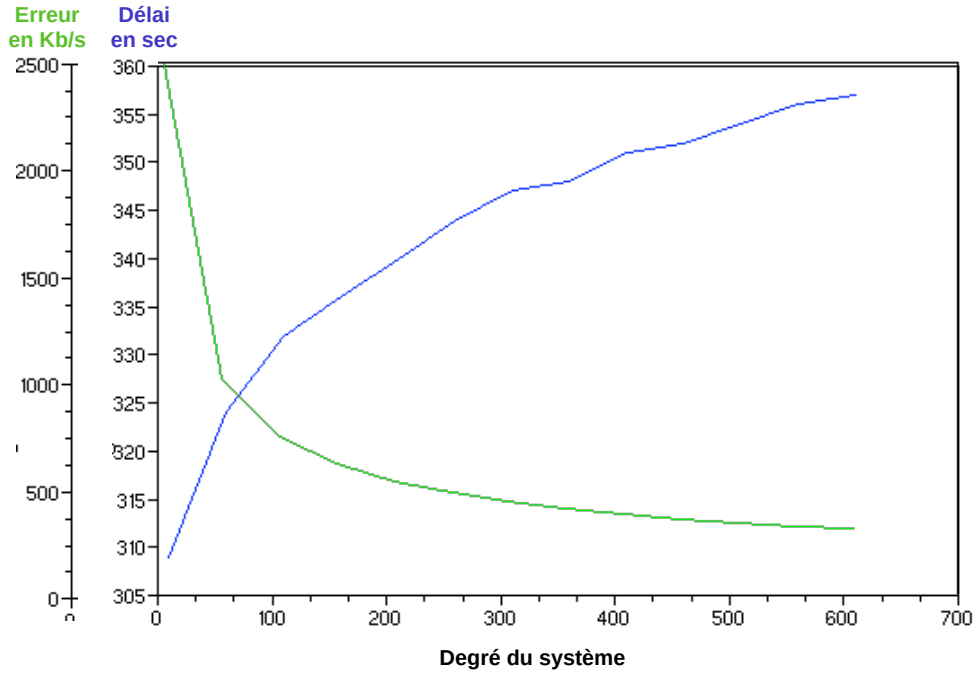

FIGURE 28: Comparaison de systèmes équivalents de dimension n

de régulateurs. Afin d'évaluer cette capacité, nous allons effectuer une comparaison entre des liaisons de même débit mais possédant un nombre N différent de régulateurs. Pour pouvoir les comparer, nous poserons les hypothèses de systèmes équivalents :

- Dans un système, tous les régulateurs sont identiques (même LLA, même HLA),
- chaque système débutera la congestion avec son CLA égal à son HLA,
- afin de pouvoir comparer les systèmes, la distance initiale entre le point de convergence et le point de fonctionnement sera la même pour tous les systèmes,
- la somme des LLA sera la même quelque soit le système,
- les valeurs de A et B seront proportionnelles au LLA,
- la liaison considérée aura un débit identique quelque soit le système.

De ces conditions on en déduit facilement les valeurs LLA, A, B en fonction du nombre de régulateurs N :

$$\begin{cases} LLA_i = \frac{\sum_{i=1}^{n} LLA_i}{n} \\ A_i = \frac{A_0}{N} \\ B_i = \frac{B_0}{N} \end{cases}$$

La valeur du CLA initiale peut également étre déduite en fonction de la distance initiale D_0, de la bande passante E et de N :

$$Distance_0^2 = \sum_{i=0}^{n} \left(CLA_i(0) - E * \frac{LLA_i}{\sum_{i=0}^{n} LLA_i} \right)^2$$

$$Distance_0^2 = n * \left(CLA_i(0) - \frac{E}{n} \right)^2$$

$$CLA_i(0) = \frac{E}{n} + \frac{Distance_0}{\sqrt{n}}$$

On peut alors comparer les performances de systèmes à degrés différents. La figure 28 montre un exemple de la variation du nombre de régulateur sur une liaison de 1Gb/s avec une somme de LLA à 500Mb/s et un facteur d'under-provisioning à 1,1. On peut observer que ce système est capable de s'adapter avec des performances comparables indépendamment du nombre de régulateurs (de 11 contrats à 100Mb/s à 550 contrats à 2Mb/s).

Etude de la dispersion des CLA_i

Lors de l'établissement des contrats, il est possible de provisionner une partie plus ou moins garantie de LLA. On peut se poser la question l'impact d'une forte disparité entre les différents contrats. Afin d'évaluer les effets, une simulation est effectuée sur une liaison 1Gb/s disposant de 50 contrats avec un taux d'under-provisioning de 50% pour le non-garanti (2Gb/s non-garanti) et un facteur Over-provisioning de 2 également pour le garanti (500Mb/s garanti). Afin de les comparer, nous établissons à l'aide d'un générateur gaussien différents CLA avec un écart-type défini. Nous avons simulé ainsi la congestion afin d'observer le délai de convergence. Nous avons répliqué ce test 50 fois afin d'obtenir le délai moyen de convergence en fonction de l'écart-type. La figure 29 représente l'évolution du délai de convergence en fonction de l'écart-type des CLA des tunnels.

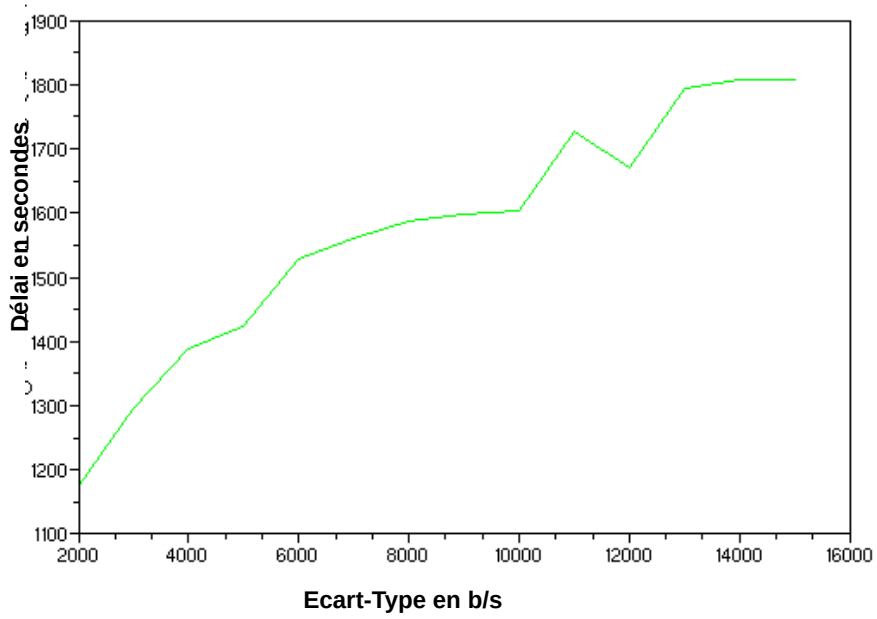

FIGURE 29: Effet de la dispersion des CLA

On remarque que plus il y a une dispersion importante du CLA plus le système met du temps à converger. Cette observation nous conforte dans le fait d'utiliser des rapports HLA/LLA fixe (chapitre 4.1.4).

4.2.4 *Propriété du au système à structure variable*

On peut constater que le système que nous avons créer fait partie de la famille des systèmes à structure variable disposant d'une surface de glissement. Dans notre cas, le systéme est définit par la surface S de glissement défini par les équations 4.16. On reprend la notation de [34] on notera ainsi $x_i = CLA_i(k)$.

$$\begin{cases} x_i(k+1) = x_i(k) - f_i\left(x_i(k)\right) & \Delta i \text{ tel que } \sum_{i=0}^{n} x_i > E \\ x_i(k+1) = x_i(k) + g_i\left(x_i(k)\right) & \Delta i \text{ tel que } \sum_{i=0}^{n} x_i < E \\ S\left(x(k)\right) = \sum_{i=0}^{n} x_i(k) - E & \text{noté } S_k \end{cases} \tag{4.16}$$

Filippov [15] démontre l'attractivité de la surface de glissement pour les systèmes à temps continu. Les travaux de Sarpturk [34] étendent cette propriété dans le cas des systèmes à temps discret, on peut alors démontrer l'attractivité d'une surface par le biais de deux inéquations qui découlent de $|S_{k+1}| < |S_k|$. La première étant la condition de

commutation (Eq. 4.17.a) et la seconde la condition de convergence (Eq. 4.17.b) :

$$\begin{cases} (S_{k+1} - S_k) * Sign(S_k) < 0 & (a) \\ (S_{k+1} + S_k) * Sign(S_k) > 0 & (b) \end{cases} \tag{4.17}$$

On démontre facilement que la première inéquation est respecté dans notre système si les fonctions f et g sont strictement positives :

$$(S_{k+1} - S_k) * Sign(S_k) < 0$$

$$\Leftrightarrow \sum_{i=0}^{n} (x_i(k+1) - x_i(k)) * Sign(\sum_{i=0}^{n} x_i(k) - E) < 0$$

si $\sum_{i=0}^{n} x_i(k) > E$	si $\sum_{i=0}^{n} x_i(k) < E$
$\Rightarrow Sign(S_k) > 0$	$\Rightarrow Sign(S_k) < 0$
$\Rightarrow (S_{k+1} - S_k) = -\sum_{i=0}^{n} f_i(x(k))$	$\Rightarrow (S_{k+1} - S_k) = \sum_{i=0}^{n} g_i(x(k))$
Or $f_i(x(k)) > 0$	Or $g_i(x(k)) > 0$
Donc $(S_{k+1} - S_k) < 0$	Donc $(S_{k+1} - S_k) > 0$

$\Rightarrow$Respect de l'inéquation 4.17a

La seconde inéquation permet de donnée les conditions supplémentaire sur f et g pour que la surface soit attractive :

$$(S_{k+1} + S_k) * Sign(S_k) > 0$$

$$\Leftrightarrow \left(\sum_{i=0}^{n} (x_i(k+1) + x_i(k)) - 2 * E \right) Sign(S_k) > 0$$

si $\sum_{i=0}^{n} x_i(k) > E$

$\Rightarrow Sign(S_k) > 0$ et

$(S_{k+1} + S_k) = \sum_{i=0}^{n} 2 * x_i(k) - f_i(x_i(k)) - 2 * E$ (cf Eq.4.16)

Pour que l'equation 4.17b soit respecté, il faut donc :

$$\Rightarrow \sum_{i=0}^{n} f_i(x_i(k)) < 2 * \left(\sum_{i=0}^{n} x_i(k) - E \right) \tag{4.18}$$

$$\text{si } \sum_{i=0}^{n} x_i(k) < E$$

$$\Rightarrow Sign(S_k) < 0 \text{ et}$$

$$\Rightarrow (S_{k+1} + S_k) = \sum_{i=0}^{n} 2 * x_i(k) + g_i(x_i(k)) - 2E$$

Ainsi pour respecté l'équation 4.17b il faudra également :

$$\Rightarrow \sum_{i=0}^{n} g_i(x(k)) < 2 * \left(E - \sum_{i=0}^{n} x_i(k) \right) \tag{4.19}$$

Par conséquence, la surface de glissement S sera attractive tant que le système respectera les équations 4.18 et 4.19.

Etude de la propriété de convergence dans le cadre du régulateur fractionnel

Dans le cadre de la modélisation précédente nous avons proposé une famille de fonction f et g. Nous pouvons ainsi établir les nouvelles conditions d'attractivité de la surface de glissement :

$$f_i(x_i(k)) = A * \frac{x_i(k)}{L_i}$$

$$\Rightarrow \sum_{i=0}^{n} f_i(x_i(k)) = A * \sum_{i=0}^{n} * \frac{x_i(k)}{L_i}$$

$$\Rightarrow A * \sum_{i=0}^{n} \frac{x_i(k)}{L_i} < 2 * \left(\sum_{i=0}^{n} x_i(k) - E \right)$$

$$\sum_{i=0}^{n} x_i(k) > E + \frac{A}{2} * \sum_{i=0}^{n} \frac{x_i(k)}{L_i}$$

$$A < \frac{2 * S_k}{\sum_{i=0}^{n} \frac{x_i(k)}{L_i}} \tag{4.20}$$

$$g_i(x_i(k)) = B * \frac{L_i}{x_i(k)}$$

$$\Rightarrow \sum_{i=0}^{n} g_i(x_i(k)) = B * \sum_{i=0}^{n} \frac{L_i}{x_i(k)}$$

$$\Rightarrow B * \sum_{i=0}^{n} \frac{L_i}{x_i(k)} < 2 * \left(E - \sum_{i=0}^{n} x_i(k) \right)$$

$$\sum_{i=0}^{n} x_i(k) < E - \frac{B}{2} * \sum_{i=0}^{n} \frac{L_i}{x_i(k)}$$

$$B < \frac{-2 * S_k}{\sum_{i=0}^{n} \frac{L_i}{x_i(k)}} \tag{4.21}$$

Du fait des termes $\sum_{i=0}^{n} \frac{L_i}{x_i(k)}$ et $\sum_{i=0}^{n} \frac{x_i(k)}{L_i}$. La limite d'attractivité est difficilement déterminable. Cependant il est possible d'encadré le résultat car tout les x_i sont supérieur à L_i et inférieur à H_i (le HLA du ième régulateur). On a alors :

$$E + \frac{A}{2} * \sum_{i=0}^{n} \frac{H_i}{L_i} > E + \frac{A}{2} * \sum_{i=0}^{n} \frac{x_i(k)}{L_i} > E + \frac{A}{2} * n \tag{4.22}$$

$$E - \frac{B}{2} * \sum_{i=0}^{n} \frac{L_i}{H_i} < E - \frac{B}{2} * \sum_{i=0}^{n} \frac{L_i}{x_i(k)} < E - \frac{B}{2} * n \tag{4.23}$$

Application numérique

Ainsi si on considère un système de 8 régulateurs passant par une liaison de 1Gb/s congestionné ou 100Mb/s ont été vendu en garantie et 200Mb/s en non-garantie pour chaque régulateur et où les constantes A et B sont positionné à 1000b/s, il est possible de déterminer des bornes aux limites d'attractivités (noté λ^+ et λ^-) de la surface grâce aux inéquations 4.22 et 4.23 :

$$1Gb/s + 4kb/s < \lambda^+ < 1Gb/s + 8kb/s$$

$$1Gb/s - 4kb/s < \lambda^- < 1Gb/s - 50b/s$$

Ce qui signifie que notre système sera attiré dans un tube autour de la surface S. Les bornes de ce tube correspondant au valeur λ^- et λ^+.

4.2.5 *Conclusion du chapitre*

Ce chapitre nous a permis de valider l'étendue des systèmes possibles. Nous avons pu montrer que pour un système donné, les performances attendues sont prédictibles et optimisables à travers le choix

des constantes et des fonctions. Les résultats obtenus pourront être utilisés pour définir des garanties contractuelles ainsi que des règles de planification et de topologies.

5

ETUDE EXPÉRIMENTALE DU SYSTÈME TDCN

5.1 PRÉSENTATION DE LA PLATE-FORME DE TESTS

L'étude théorique précédente a mis en évidence de multiples intérêts à mettre en oeuvre notre système. Cependant, l'ensemble des résultats obtenus repose sur une modélisation théorique. Lors de l'établissement du modèle, nous avons dû poser des hypothèses et effectuer des simplifications. Dans le cadre de la validation du modèle, nous avons réalisé une maquette afin de vérifier la pertinence des hypothèses prises. D'autre part son second objectif est d'apporter des informations sur la qualité de service qui n'a pas été modélisée.

5.1.1 *La plate-forme de tests*

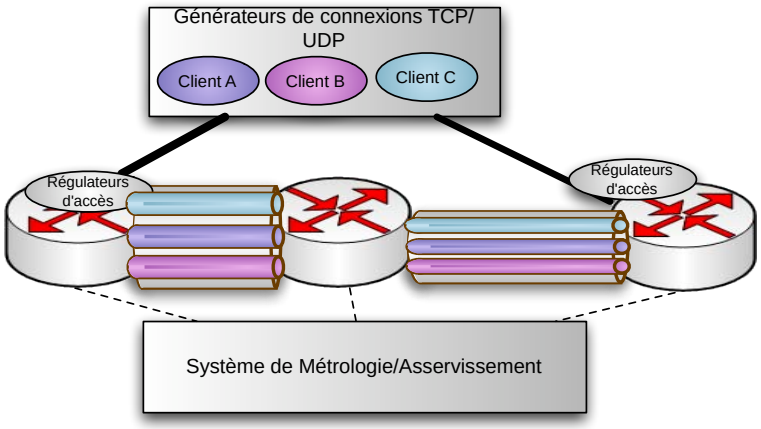

FIGURE 30: Schéma du testbed

L'objectif de la plate-forme est de retrouver l'ensemble minimal des composants permettant la mise en oeuvre du système TDCN. L'environnement d'un WAN est émulé afin de retrouver les comportements

TDCN identique à celui que l'on aurait dans le cadre d'un déploiement de la technologie. On retrouvera également les générateurs nécessaires à l'établissement des flux représentatifs des agrégats réels. Enfin la maquette reposera également sur la mise en place d'un système de métrologie le plus complet possible.

Afin d'observer le système en phase de congestion, la maquette se compose de trois routeurs : un point de congestion est alors créé sur le routeur central afin de simuler une saturation de coeur de réseau. Afin de réaliser cette congestion, nous effectuons une limitation du débit d'émission sur l'interface de sortie du routeur de coeur (limite à 80Mb/s).

5.1.2 *Création de l'environnement de tests*

le générateur d'agrégats

FIGURE 31: Générateur de session NetworkTester

Le système TDCN est conçu pour être déployé dans un réseau de transport mutualisé. Afin de retrouver les comportements de flux clients et plus particulièrement les propriétés d'élasticité des flux, nous utilisons un générateur industriel de sessions TCP.

L'équipement utilisé est un générateur NetworkTester de la société Agilent. Celui-ci permet d'émuler des applications basées entre autre sur TCP. Nous avons plus particulièrement utilisé ses capacités à émuler des sessions clients/serveurs FTP. Notre but était uniquement de profiter des capacités TCP et de contrôler le nombre de connexions TCP simultanées. Ce nombre de session étant l'un des principaux paramètres influençant le partage de la bande passante, nous pouvons ainsi étudier le comportement du système TDCN dans différentes conditions de régimes.

L'émulateur WAN

L'environnement d'un réseau transport mutualisé est bien différent de celui retrouvé dans un laboratoire. Afin de reproduire les effets de cet environnement, nous utilisons un émulateur. Celui-ci a pour principal objectif de recréer une latence dans le réseau afin que les comportements des sessions TCP soient plus représentatifs de la réalité. Nous avons donc utilisé un serveur Windows dédié disposant du programme NetDisturb edité par Omnicor afin de réhausser la latence du réseau. Nos tests ont été effectués avec un délai réseau de 20ms.

5.1.3 *Implémentation du système TDCN*

Les routeurs

FIGURE 32: Routeur 7450 ESS1

La réalisation de notre maquette a été effectuée avec des routeurs de coeurs Alcatel-Lucent. Le choix s'est porté sur l'utilisation de routeurs industriel dont nous connaissions parfaitement le dimensionnement de la capacité de commutation ainsi que celle de ses différentes cartes porteuses. Ces routeurs disposent également d'une stack MPLS complète incluant la gestion des VPN. Ils possèdent également la capacité pour les ports d'accès au VPN de gérer près de 16000 files d'attente différentes et autant de QoS d'accès. Ce qui le rend particulièrement intéressant pour notre système qui prévoit l'utilisation de régulateur dédié à chaque tunnel.

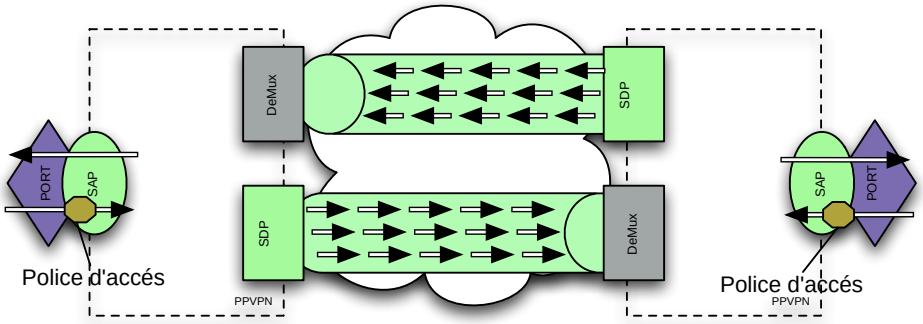

FIGURE 33: Architecture Générique d'un VPN

Description de l'architecture VPN

Dans le cadre de leur utilisation normale, les routeurs cloisonnent leur clients dans des services VPN différents. Chaque service VPN (fig. 33) de ces routeurs est composé de plusieurs éléments :

- **Service Access Point (SAP) :** Entité rattachée à un port ou à un VLAN d'accès. On peut lui appliquer une politique particulière de qualité de service.
- **Service Distribution Point (SDP) :**Entrée du tunnel MPLS . Il a pour charge d'émettre les paquets dans le réseau. Il peut posséder des règles de filtrage spécifique en fonction de la nature des flux.
- **Demux :** Il récupère les paquets des tunnels MPLS. Traite les informations MPLS et les retourne dans le service PPVPN en correspondance.
- **Sap-ingress policy :** Politique d'accès des SAP. Réserve un espace mémoire pour le SAP. Gère la colorie par le biais d'un TRTCM [18]. Le même mécanisme existe également en sortie de SAP (Sap-egress policy).

Notre Système TDCN utilisera cette architecture, un service TDCN sera un VPN point à point disposant de politiques d'accès spécifiques.

L'emulateur TDCN

Etant donné que les routeurs ne supportent pas nativement notre nouveau système expérimental, nous avons été contraint d'externaliser sur un serveur Linux les briques nécessaires à la mise en oeuvre de notre système TDCN. Trois composants sont ajoutés :

– Des contrôleurs de congestion dans les routeurs de coeur chargés de détecter une éventuelle congestion.
– Un protocole de signalisation des congestions.
– Des régulateurs d'accès dynamiques prenant en compte les informations transmises par le protocole de signalisation.

Pour ce faire, nous avons utilisé des automates logiciels qui interrogent et contrôlent les routeurs sur un réseau de management dédié. Ceux-ci récupèrent les informations et modifient le comportement du routeur par le biais du protocole SNMP. Ce fonctionnement entraîne toutefois une moins grande réactivité. Nous verrons par la suite quelle peut être l'impact de cette méthode sur les résultats.

Les contrôleurs de congestion

Le contrôleur de congestion est situé au niveau de chaque port de coeur de réseau. Il contrôle en permanence l'état des buffeurs. Il retransmet ensuite l'information de cet état toutes les secondes au protocole de transport d'information.

En théorie, il faut contrôler l'état des buffeurs de sortie de l'interface qui se rempliront dans le cadre de la congestion de la liaison mais également les buffeurs d'entrées des interfaces qui eux se rempliront lorsqu'une congestion du coeur de commutation apparaît. Dans la pratique, la capacité de commutation maximum de nos routeurs n'est pas sous dimensionnée, et cette congestion ne peut pas apparaître.

Implémentation des DAR

Afin de réaliser les DARs, nous avons utilisé des TRTCMs [18] inclus au niveau des politiques d'entrée des SAP des routeurs. Nous avons ensuite piloté via le protocole SNMP la valeur du PIR de celui-ci. Ainsi, le CLA de notre système correspond à la valeur courante du PIR du TRTCM. Le LLA est lui représenté par le CIR du régulateur.

En théorie, ces TRTCM devraient être appliquées au niveau de policy appliqués sur les SDP. Cependant, cette configuration n'est pas faisable matériellement d'où sa mise en oeuvre sur le SAP. Cette modification n'entraîne aucune conséquence dans le cadre d'un VPN point à point. Par contre, elle empêcherait l'utilisation de VPN multipoints. Dans le cadre de l'industrialisation du procédé, il conviendrait de disposer les mécanismes sur le SDP.

Le protocole de transport d'information

L'ajout ou la modification d'un protocole au niveau des routeurs implique nécessairement un recodage de celui-ci. Aussi nous avons choisi d'émuler le comportement du protocole. Nous avons co-localisé les programmes des contrôleurs de congestion et des DARs sur le serveur GNU/Linux avec une transmission d'informations par le biais de variables partagées. Pour autant, les divers codes sont désynchronisés tout comme ils le seraient si le code était inclus dans leurs routeurs respectifs.

5.1.4 *Défaut de l'externalisation du code*

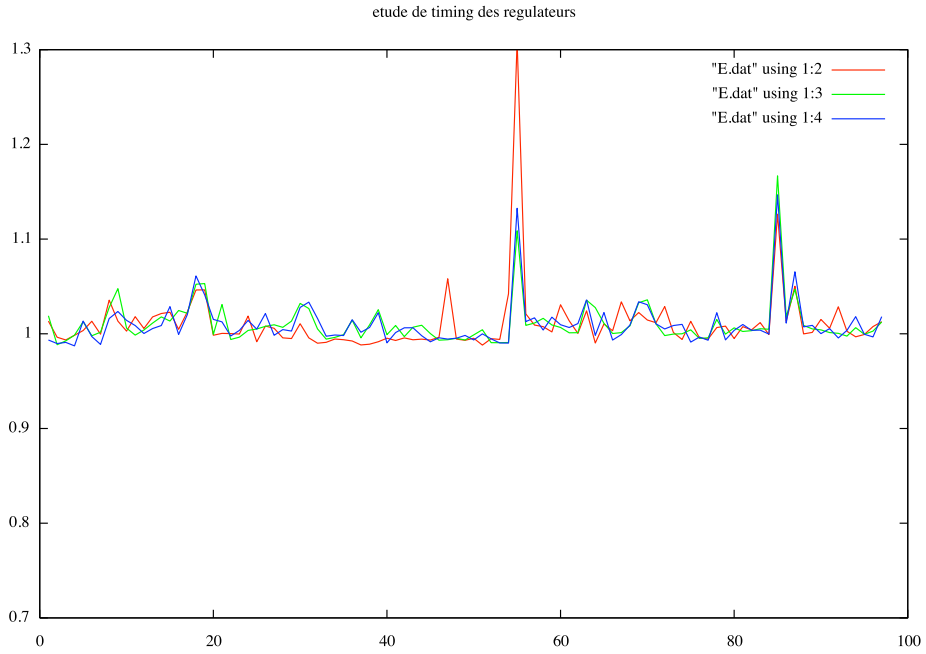

FIGURE 34: Variation du temps entre les passes algorithmiques

Du fait de l'utilisation d'un OS non temps réel et également à cause des délais sur le réseau de management, les DAR n'agissent pas toutes les secondes précisément. La figure 34 montrent un exemple de variation de délai entre 2 passes de l'algorithme. Ces courbes devraient

en théorie être des droites constantes à 1 seconde. Cette variation serait largement atténuée si le code était inclu directement dans les routeurs.

5.1.5 *Mesures de la plate-forme de tests*

Métrique du routeur

Le routeur dispose pour chacun des ports des informations suivantes : compteurs d'octets, de paquets et de drops. Ces compteurs sont rafraîchis toutes les secondes. Mais, il est également possible de connaître l'état des buffeurs des ports d'accès et de coeurs (fig. 35) pour chaque file d'attente.

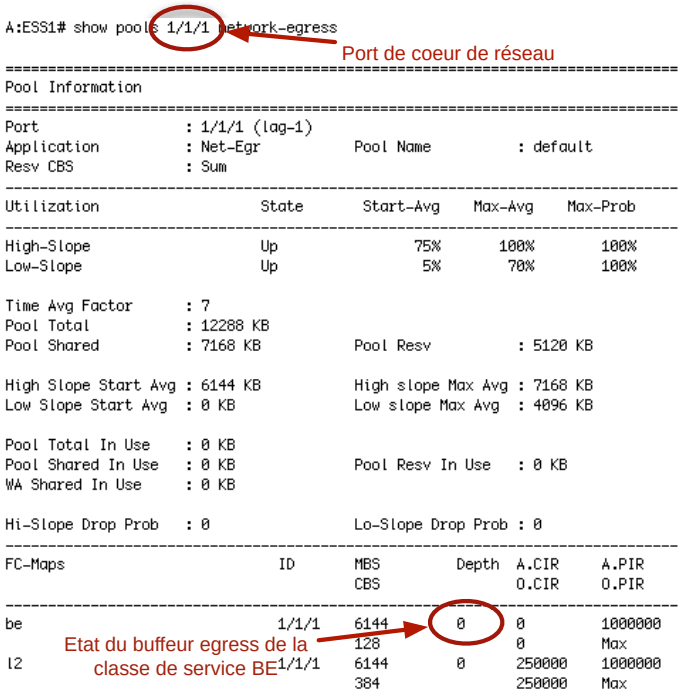

FIGURE 35: Commande show pool sur un routeur ESS1

Afin de suivre l'évolution de ces informations, des automates d'interrogations ont été réalisés pour nous permettre de mesurer l'évolution de ces valeurs. Contrairement aux autres compteurs, l'interrogation de l'état des buffers prend entre 400 et 600ms et donne une information sur un état instantané.

La figure 36 est un exemple d'évolution d'un buffeur suite à l'apparition d'une congestion de coeur.

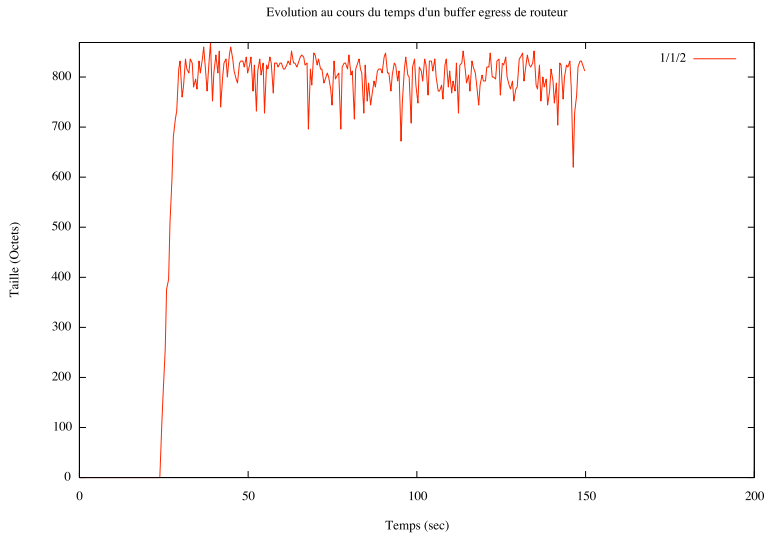

FIGURE 36: Evolution d'un buffeur de sortie sur un port en coeur d'un routeur disposant d'une liaison congestionnée

Le générateur de débit

Afin de mesurer les performances des systèmes mis en place, nous avons utilisé un générateur de trafic N2X de la marque Agilent. Cet équipement génère des paquets IP auxquels il ajoute un timestamps d'horloge dans la payload. A l'aide d'un second port, il récupère ces paquets et peut ainsi mesurer la latence, la gigue, le taux de pertes, les désordonnancements, le tout avec une précision de 7 picosecondes.

5.2 PLATEFORME DE TEST DU SYSTÈME TDCN

Le chapitre précédent nous a permis d'exposer le fonctionnement de la plate-forme de test mettant en oeuvre l'approche TDCN. Dans ce chapitre nous allons utiliser cette plate-forme pour valider la modélisation effectuée et mesurer les performances réseaux du nouveau système.

5.2.1 *Comparaison entre les résultats de la maquette et de la modélisation*

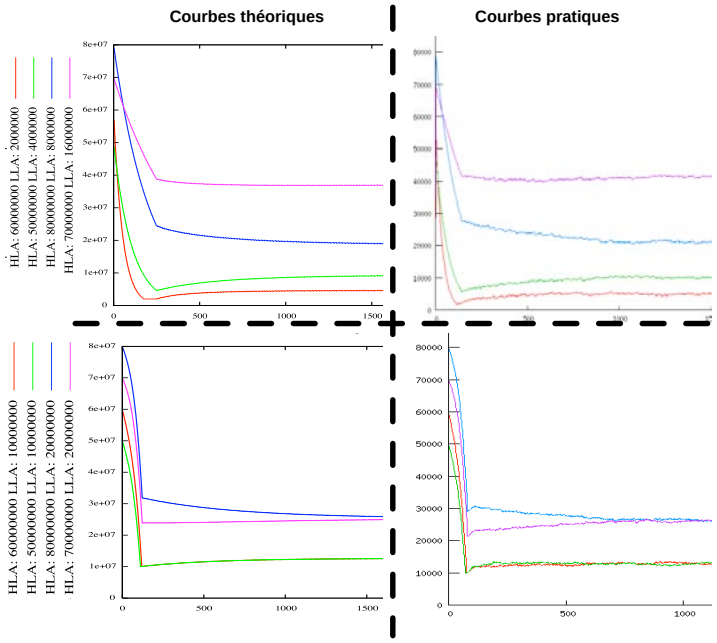

FIGURE 37: Comparaison entre les résultats sur la maquette et dans la modélisation

L'établissement de la modélisation des chapitres précédents a été réalisé sur la base des principes théoriques issue des mécanismes mis en oeuvre. La complexité des flux réels et les simplifications que nous avons dû opérer pour établir le modèle, nous amènent à vérifier l'adéquation entre notre approche théorique et une mise en oeuvre pratique.

Pour ce faire, nous avons étudié l'évolution des CLA des régulateurs du système TDCN dans le cadre de notre plate-forme de test ; puis nous allons comparer ces résultats face à l'estimation de cette évolution via la modélisation. Dans une première approche, nous ne nous focaliserons pas sur les débits réels présents dans chaque VPN mais uniquement sur l'évolution des régulateurs. Nous positionnerons simplement suffisamment de connexions TCP en concurrence pour initier une congestion.

La figure 37 compare les résultats théoriques et pratiques dans le cadre de deux systèmes paramétrés différemment. On constate sur le

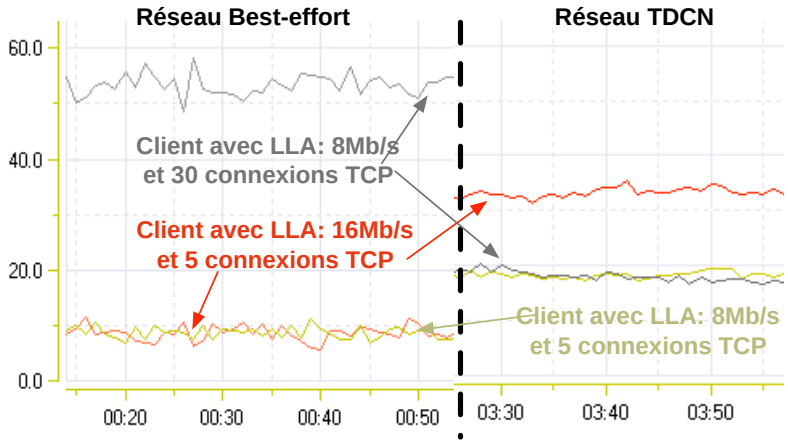

FIGURE 38: Evolution du trafic client dans un réseau BE et TDCN

graphique que les résultats de la modélisation, bien que plus lissés que ceux mesurés en pratique, donnent une bonne estimation de l'évolution réelle des CLA.

Ce lissage s'explique facilement par le fait que la modélisation considère à tort une action simultanée de tous les régulateurs. L'effet de la désynchronisation se traduit par une variation plus importante des CLA. Par contre, le délai d'équilibrage est légèrement surévalué dans le cadre de la modélisation.

5.2.2 *Étude des flux*

Fonctionnement en régime établi

Nous avons précédemment validé le bon comportement des régulateurs. Nous allons maintenant nous attacher à vérifier l'évolution des flux réels inclus dans ces tunnels. Nous utiliserons pour ce faire le système de métrologie défini dans le chapitre précédent pour étudier l'évolution des débits clients.

L'exemple présenté dans la figure 38 a été réalisé sur la base de trois tunnels TDCN dans lesquels chaque client utilise des connexions TCP simultanées. Dans l'exemple retenu, le premier client a été configuré pour exécuter 30 connexions TCP simultanées. Les deux autres clients ont été configurés pour exécuter chacun cinq connexions TCP.

Dans une première étape, Nous avons mesuré l'utilisation de la bande passante pour chaque client dans un réseau classique Best-effort

sans mis en place de notre système. Dans un second temps, nous avons réitéré les mêmes mesures avec notre réseau TDCN.

On peut remarquer que dans le réseau classique le partage est effectué au prorata du nombre de connexions TCP. De nombreuses études donnent aussi le même type de résultat [25], et décrivent les mécanismes de partage du réseau. Ils soulignent en particulier la dépendance entre la notion de partage et la quantité de connexions TCP.

Cette situation est particulièrement gênante pour un opérateur, puisque des clients disposant de contrats identiques ne pourront pas revendiquer la même bande passante. Un client malveillant pourrait dans certains cas de figure être tenté d'augmenter virtuellement son nombre de connexions TCP afin de récupérer plus de bande passante au détriment des autres clients.

Le système TDCN a justement été conçu pour empêcher la nature des flux d'influer sur le partage de la bande passante. La figure 38 met bien en évidence cette nouvelle indépendance aux flux. En outre, la bande passante est partagée proportionnellement au paramétrage des LLA des régulateurs. La figure 39 montre que tous les clients ont des taux de LLA à CLA identiques, ce qui signifie que la bande passante supplémentaire a été équitablement répartie entre les différents clients, indépendamment du nombre de connexions TCP.

Dans la pratique, les flux clients sont contraints individuellement par les limitations données par les régulateurs DAR. Il n'existe alors aucune possibilité pour les clients de se substituer à cette limite.

Etude en phase d'équilibrage

Nous avons vu précédemment par le biais de la modélisation que notre système possède une phase transitoire au niveau des régulateurs. Dans cette partie, nous focaliserons notre attention sur l'évolution des trafics clients durant cette période. Afin de contrôler l'évolution lors de cette phase, nous allons initier une congestion dans le coeur de réseau en établissant simultanément les trafics clients. La configuration mise en oeuvre est reprise du test effectué dans le chapitre précédent. Au cours de l'essai, nous avons étudié le débit utilisé pour chaque client (figure 40). Dans le même temps, l'évolution des valeurs CLA a été enregistrée pour chaque régulateur. (Figure 41).

A l'initialisation de la congestion, nous observons que tous les trafics se partagent la bande passante de la liaison. Ce partage de

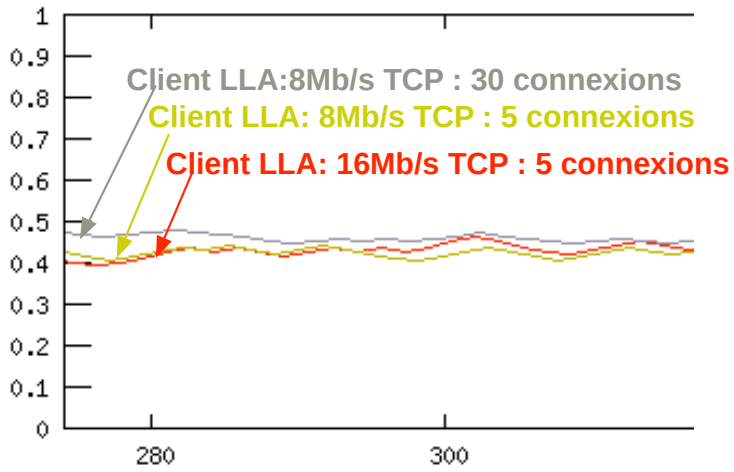

FIGURE 39: Evolution du rapport LLA/CLA des régulateurs dans le réseau TDCN

bande passante est similaire à celui qui aurait eu lieu dans un réseau classique best-effort(voir figure 38).

Par la suite les régulateurs vont agir face à cette congestion en durcissant leur politique d'accès. Le trafic, précédemment le plus favorisé par le partage best-effort, sera le premier impacté et devra réduire son utilisation de la bande passante. Dans le même temps, la bande passante libérée sera réallouée entre les autres clients qui n'avaient pas atteint leurs limites.

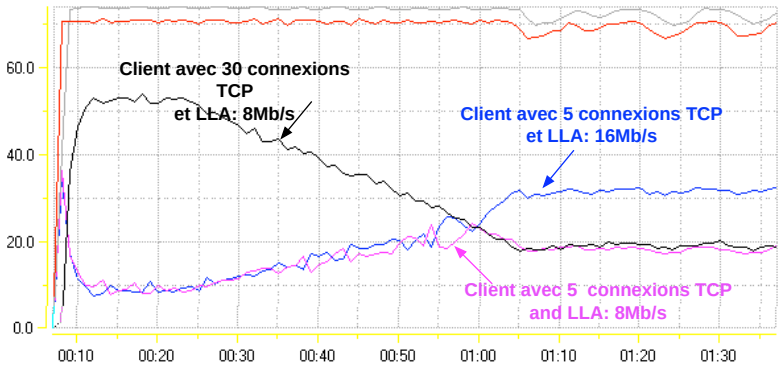

FIGURE 40: Evolution du trafic client en phase transitoire

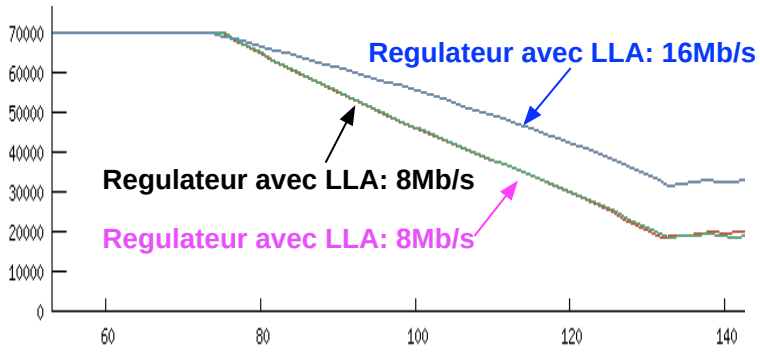

FIGURE 41: Evolution du CLA des clients en phase transitoire

On constate également que les flux TCP subissant le durcissement de politique des DAR ne rompent pas leurs connexions et suivent la pente donnée par le régulateur. Ce comportement est rendu possible car l'échelle de temps de la variation des régulateurs est significativement supérieure à celle des connexions TCP (facteur 100).

5.2.3 *Étude des performances globales*

Evaluation des performances par client

Les résultats précédents ont mis en évidence la capacité de notre système à fournir une équité contractuelle lors d'une période de congestion après une phase transitoire. Dans cette partie nous allons étudier les performances qui découlent de ce fonctionnement dans un cas extrême ou la garantie contractuelle est inversement proportionnelle au nombre de connexions des flux clients.

L'expérimentation suivante se compose de sept tunnels TDCN mis en concurrence, chaque client disposant d'un nombre différent de connexions TCP. Chaque connexion TCP effectue le téléchargement d'un fichier de 5 Mo, une fois la connexion terminée une nouvelle connexion est rétablie (18 fois pour chaque connexion).

Nous mesurons le temps pour finaliser l'ensemble des téléchargements pour chaque client et nous calculons ainsi le débit moyen obtenu en pratique par client. Nous effectuons l'expérience dans le cadre d'un réseau classique BE puis dans le cadre de notre proposition d'architecture.

Les résultats obtenus sont synthétisés dans le tableau 1. Ce test dure quelques minutes. Afin d'étudier l'évolution à plus long terme de

TABLE 1: Performance par client pour 18 téléchargements successifs

Client	LLA	Connexions TCP	Durée TDCN	Débit TDCN	Débit BE
1	14Mb/s	1	101 s	7,1Mb/s	2,4Mb/s
2	12Mb/s	2	132 s	10,9Mb/s	4,7Mb/s
3	10Mb/s	3	170 s	12,7Mb/s	7,1Mb/s
4	8Mb/s	4	210 s	13,7Mb/s	9,4Mb/s
5	6Mb/s	5	250 s	14,4Mb/s	11,8Mb/s
6	4Mb/s	6	280 s	15,4Mb/s	14,2Mb/s
7	2Mb/s	7	345 s	14,6Mb/s	16,5Mb/s

TABLE 2: Performance par client pour 200 téléchargements successifs

Client	LLA	Connexions TCP	Durée TDCN	Débit TDCN	Débit BE
1	14Mb/s	1	555 s	14,4Mb/s	2,4Mb/s
2	12Mb/s	2	1025 s	15,6Mb/s	4,8Mb/s
3	10Mb/s	3	1485 s	16,2Mb/s	7,2Mb/s
4	8Mb/s	4	2005 s	16,0Mb/s	9,4Mb/s
5	6Mb/s	5	2480 s	16,1Mb/s	12,1Mb/s
6	4Mb/s	6	3185 s	15,1Mb/s	14,5Mb/s
7	2Mb/s	7	3485 s	16,1Mb/s	16,9Mb/s

la performance, un deuxième test a été réalisé, fondé sur le même paramétrage excepté pour le nombre de téléchargements qu'on réitère 200 fois. L'expérience dure alors plus d'une heure.

Nous pouvons observer dans le tableau 1, que tous les clients ne sont pas en mesure de respecter le contrat minimum LLA. Ce phénomène s'explique par le fait que pendant la phase de transition, le partage de la bande passante est effectué au pro-rata des connexions TCP. Il

s'agit donc d'un partage diamétralement opposé à celui souhaité dans ce cas particulier. Toutefois, même dans ce cas exagérément négatif, les résultats sont bien meilleurs que pour le réseau Best-effort (le client dispose d'un débit trois fois supérieur à celui qu'il aurait eu en Best-effort).

On peut remarquer également que, même sans tenir compte des contraintes liée au LLA, six des sept clients ont un débit moyen supérieur à celui qu'ils auraient obtenu dans un réseau Best Effort. Ce phénomène s'explique par le fait qu'une fois le téléchargement d'un client terminé, la bande passante libérée est redistribuée aux clients qui n'ont pas encore fini leurs téléchargements. Ainsi on peut considérer que, si les demandes de bande passante d'un client sont en quantité finie, le système n'obère pas le rendement du réseau. Son fonctionnement se traduit par un réordonnancement dans le temps des demandes clientes.

Ce test met en avant un certain nombre d'avantages de notre système :

- Tout d'abord, même dans le cas le plus négatif où le nombre de connexions TCP est inversement proportionnel aux contrats du client, le partage reste extrêmement avantageux pour les contrats les plus importants.
- Si on considère que les demandes des clients ne sont pas infinies, les contrats les plus bas ne sont pas trop impactés.

Mesures des performances globales du système

Nous allons utiliser les résultats des tests précédents afin d'évaluer la performance globale de notre réseau. En fait, un réseau parfait devrait être en mesure de transporter tous les trafics à la bande passante du goulot d'étranglement. Le délai le plus court pour effectuer l'ensemble des téléchargements pourrait être déterminé par la formule suivante :

$$Delai_{Parfait} = \frac{18 * (5MB/s * 8) * 28Connexions}{76Mb/s} = 265s \qquad (5.1)$$

Ainsi la performance globale du système peut être mesurée par le rapport entre le délai pour terminer l'ensemble des téléchargements et le délai du système parfait :

$$R_{system} = \frac{Delai_{Mesure}}{Delai_{Parfait}} \qquad (5.2)$$

Nous observons que le nouveau système est moins efficace qu'un réseau Best-effort principalement pour le premier jeu de données. La

cause principale de cette perte d'efficacité repose sur le délai nécessaire pour redistribuer la bande passante libérée par un client. Le premier test met en exergue ce phénomène. Nous pouvons voir un exemple pratique de cet évènement sur la figure 42.

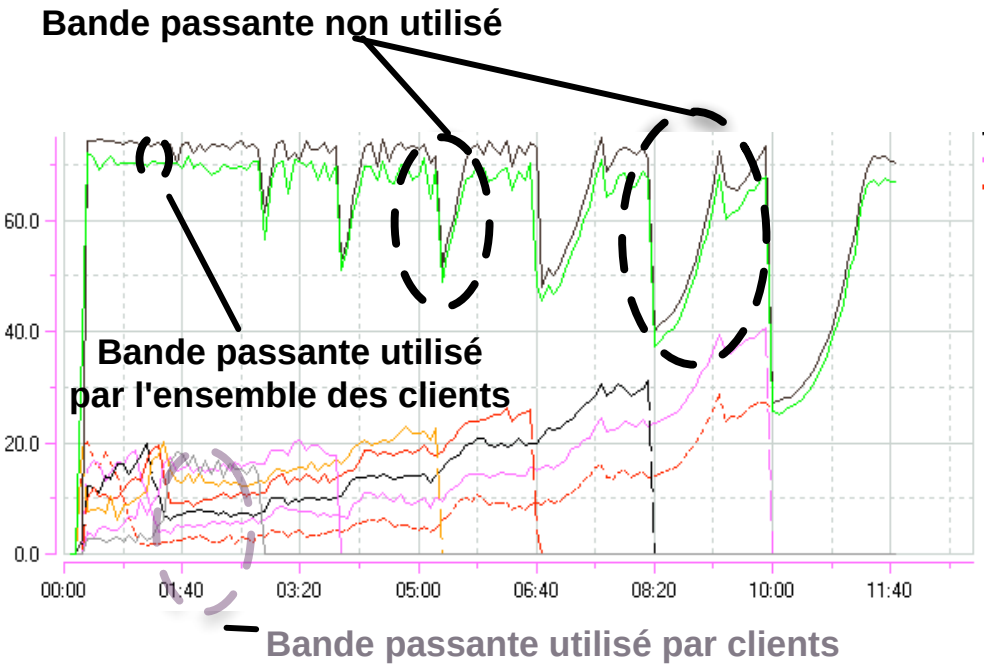

FIGURE 42: Exemple de délai pour réallouer la bande passante

Dans cette exemple, les clients libèrent soudainement l'ensemble de leur trafic entraînant ainsi une phase temporaire où l'ensemble de la bande passante n'est pas utilisée dans le coeur du réseau. Ce phénomène dans le cas de flux agrégés n'a que très peu de chance de ce produire. Notre système est donc mieux conçu dans le cadre de partage entre flux agrégés.

5.2.4 *Etude de la qualité de service*

Nos études ont mis en évidence les capacités en terme de gestion de la bande passante du système. Cependant il reste à analyser les performances qualitatives des flux transportés, plus particulièrement

TABLE 3: Estimation de la performance global du système

Systéme	Rendements
Premier jeu Best-Effort	87%
Premier jeu TDCN	77%
Second jeu Best-Effort	89%
Second jeu TDCN	85%

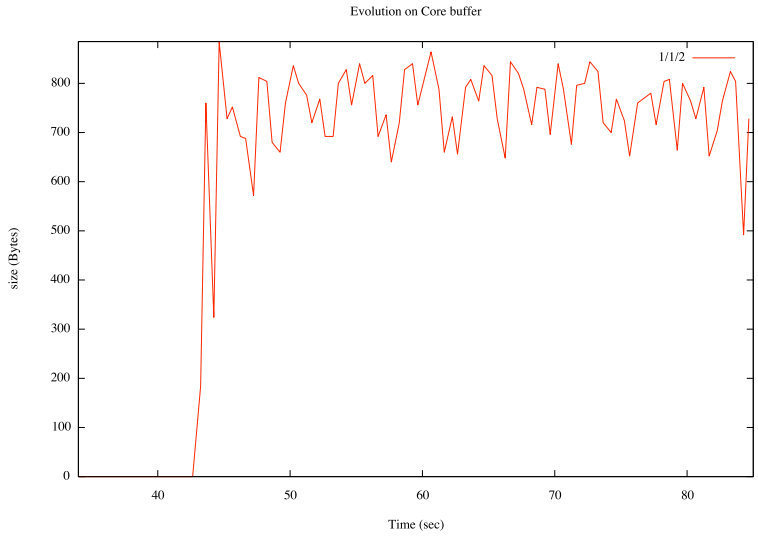

FIGURE 43: Evolution du buffer d'un réseau Best-effort en congestion

pendant la phase de congestion. Ces performances sont alors directement liées à l'évolution des buffers des routeurs.

Étude de l'évolution des buffers

Dans un réseau Best Effort standard, la période de congestion entraîne une bufferisation au niveau du goulot d'étranglement (figure 43). Le niveau d'encombrement du buffer dépend alors de différents paramètres dont le profil des flux clients (nombres de connexions TCP, débit UDP,...), mais également de certains paramètres des routeurs par exemple le mécanisme RED [4],[26].

Dans notre réseau TDCN on peut noter que le profil de bufferisation (figure 44)est très différent. Dans la phase transitoire, on retrouve un

profil analogue à celui qu'on observerait sur un réseau Best Effort avec toutefois une différence notoire : les flux les plus agressifs subissent une deuxième bufferisation (moins importante) dans le buffer d'accès du régulateur. Cette seconde bufferisation est une conséquence directe du durcissement de la limite CLA.

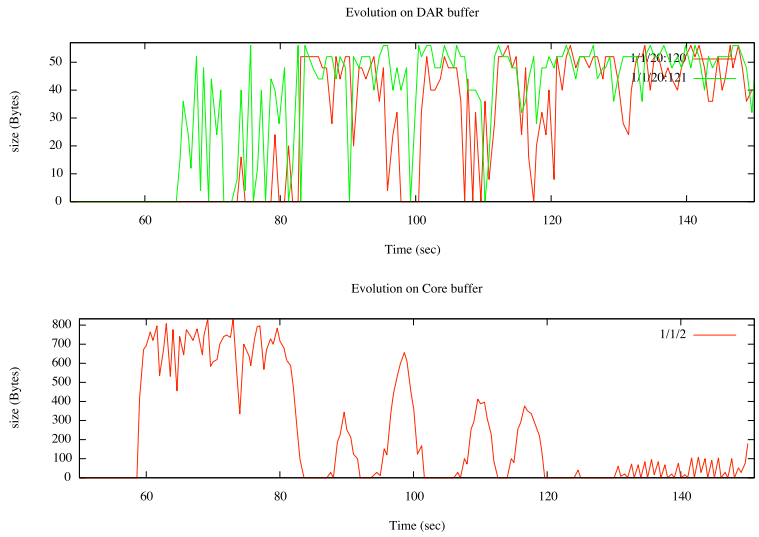

FIGURE 44: Evolution du buffers dans un réseau TDCN

Par contre, une fois la période d'équilibrage atteinte, le profil de bufferisation change complètement. Vu du flux TCP, la bufferisation bascule constamment entre une présence en coeur de réseau et une présence au niveau du régulateur du routeur d'accès. Il s'agit d'un attribut très intéressant parce que la congestion présente au niveau du buffer d'accès est bien moins conséquente qu'en coeur. En fait, la congestion présente en coeur est due à l'aggressivité de l'ensemble des flux TCP passant sur le lien en congestion, par contre en accès seules les connexions TCP du client participent à l'évolution du buffer. Ce phénomène se traduit directement sur la latence des flux clients qui est inférieure à celle qu'elle aurait eu dans un réseau Best-effort. Notre système transforme ainsi une congestion de K clients en K congestions réparties sur l'ensemble du réseau.

Effet du système sur les trafics UDP

Jusqu'à présent nos études se sont focalisées sur des flux TCP. Dans cette partie, nous contrôlerons l'impact sur les trafics UDP. L'évolution de la charge des buffers dans le cadre de notre système, entraîne différents effets sur le trafic UDP.

Nous avons réalisé un test en injectant un trafic UDP instrumenté avec le générateur N2X. Ce test a été effectué avec un trafic UDP en Constant Bit Rate (3750 fps avec des paquets de 100 octets) injecté dans chaque tunnel client. La figure 45 montre l'évolution de la latence et des pertes de paquets de ces trafics. Le système est composé de trois clients :

– le premier dispose de 30 connexions TCP,
– les deux autres disposent de 5 connexions chacun.

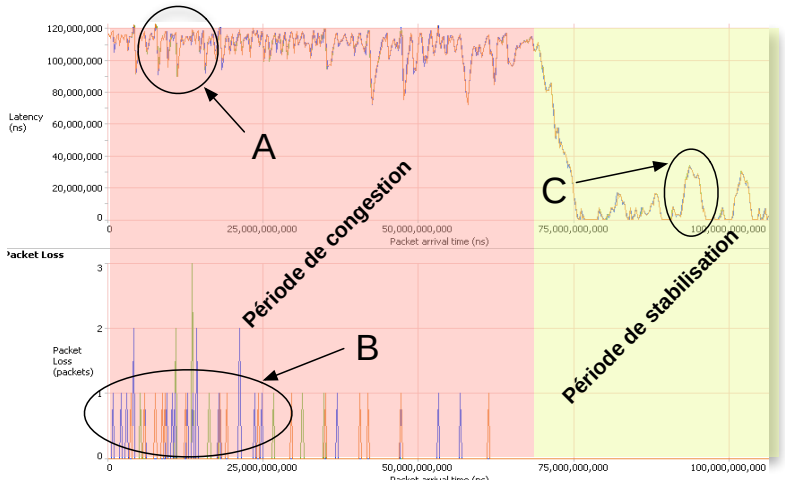

FIGURE 45: Evolution de la latence et des pertes de paquets des flux UDP dans un réseau TDCN

Nous observons sur la figure 45 trois phénomènes remarquables :

– **point A :** Pendant la phase transitoire du système, la latence des flux UDP est la même que celle qui aurait été obtenue dans un réseau Best Effort. Dans notre exemple, cette latence est de 120 ms.
– **point B :** Dans le même temps, on observe des pertes sur les paquets UDP. Ces pertes sont dues d'une part au mécanisme RED mis en place dans le coeur de réseau et d'autre part au

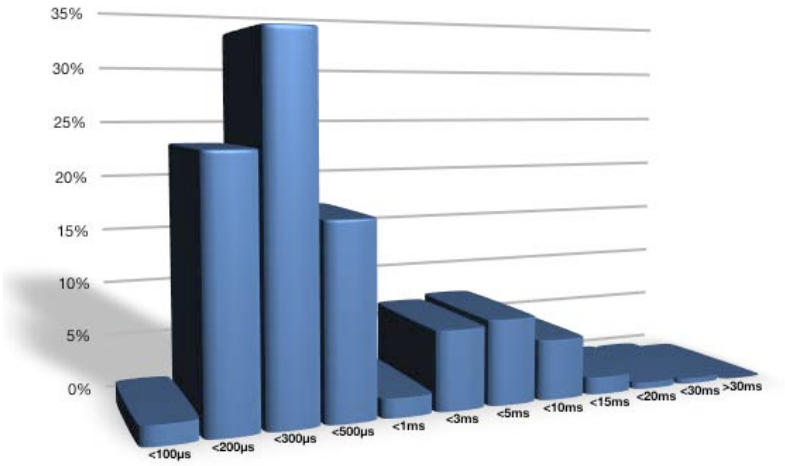

FIGURE 46: Distribution de la latence pour les flux UDP dans un réseau TDCN

durcissement de la politique d'accès des régulateurs DAR. A noter que seuls les flux les plus agressifs subissent cette deuxième cause (ici le client qui dispose de 30 connexions TCP).

– **point C :** Durant la phase de stabilisation, du fait du passage temporaire en congestion de coeur, la latence des flux UDP augmente temporairement. On peut remarquer que cette augmentation de latence se produit périodiquement et reste inférieure à une bufferisation observée en congestion classique de coeur (25ms contre 120ms). Hors de cette période, la latence est similaire à celle d'un réseau non-congestionné.

Finalement, une fois la période de stabilisation atteinte, notre système dispose d'une distribution de latence atypique car basée sur deux bufferisations successives. En effet, l'oscillation entre la congestion de coeur et d'accès entraîne une variation de la latence des paquets. La figure 46 montre cette distribution de la latence. On peut noter que cette gigue n'entraîne aucun désordonnancement des paquets. Il est par ailleurs possible de modifier cette distribution en jouant sur les paramètres de montée et de descente des régulateurs. Par exemple, si le paramètre de montée est beaucoup plus lent que celui de descente, le système restera en congestion d'accès plus longtemps qu'en congestion de coeur et donc nous aurons une latence globale plus faible. et en tout cas inférieur à celle 'une congestion de coeur.

5.2.5 *Conclusion du chapitre*

La plate-forme de test nous a permis de confronter les résultats théoriques avec les mesures pratiques. Les doutes sur d'éventuelles interactions entre la régulation effectuée par TCP et la sur-régulation TDCN ont ainsi pu être levés. De plus, nous avons mis en évidence de nouvelles propriétés intéressantes de notre système comme par exemple, le nouvel état de quasi-congestion créé par le mécanisme TDCN qui apporte une utilisation presque maximale de la bande passante tout en conservant une latence très faible.

CONCLUSION

Dans ce mémoire de thèse, nous avons présentés nos travaux orientés sur la problématique de mutualisation de ressources dans le cadre de réseau d'opérateurs de transport mutualisé. Cette problématique résulte d'une volonté antagoniste de fournir d'une part un maximum de bande passante à chaque client et d'autre part de disposer d'une qualité et d'une disponibilité maximales sur l'ensemble du réseau.

En effet, la demande de bande passante est en constante augmentation. De plus l'avènement des accès FTTH (Fiber To The Home) pour l'Internet, la 3G pour la téléphonie mobile, et de la vidéo haute définition entraîne une pression supplémentaire sur les coûts. Les opérateurs cherchent alors à en limiter les répercutions sur les utilisateurs en optimisant leurs infrastructures. Les travaux de recherche se focalisent alors sur l'optimisation des coeurs de réseaux afin d'en limiter le surdimensionnement habituel.

Dans le cadre d'une recherche de réponse à cette problématique, nous avons tout d'abord montré les possibilités offertes par les méthodes de provisioning. Celles-ci s'attachent à modifier les chemins des flux pour maximiser l'utilisation du coeur de réseau. Malgré tout ces approches reposent sur une vision macroscopique des flux. Dans un second temps, nous avons étudié les approches historiques gérant le partage de la bande passante entre les session TCP. Celles-ci se focalisent sur l'obtention d'une équité protocolaire entre les diverses sessions. On peut parler d'une vision microscopique des flux.

A partir de ces études, nous avons mis en évidence la nécessité de transmettre la gestion de l'équité à des protocoles internes à l'opérateur de l'infrastructure pour permettre une gestion plus approfondie du provisioning réseau.

Nous avons ainsi établi une nouvelle approche appelée TDCN qui permet de relier ces deux approches. Notre proposition a pour objectif de transmettre le partage de la bande passante au niveau des paramètres des tunnels donnant ainsi le contrôle de la congestion à l'opérateur de l'infrastructure. Le design modulaire du système permet de nombreux paramétrages en fonction de la politique de chaque opérateur.

Par le biais d'une modélisation, nous avons démontré le fonctionnement du système. Nous avons proposé également une méthode

permettant d'évaluer l'efficacité du mécanisme en fonction du paramétrage du système. Nous avons établi ainsi un sous-système basé sur la terminologie des contrats actuels d'opérateurs permettant une industrialisation.

Une plate-forme de test a permis de valider notre modèle en démontrant le bon fonctionnement du système proposé dans un cadre réel et en mettant en évidence les intérêts majeurs en terme de qualité de service.

Le système a été étudié pour fournir une garantie de débit pour un service transitant au niveau d'un AS. Des perspectives de recherche apparaissent alors : il serait intéressant d'étudier la possibilité de fournir une qualité de service de bout en bout en profitant des garanties locales, à l'image des travaux qui furent réalisés pour pouvoir fournir une qualité Intserv en traversant des domaines Diffserv ([27]).

D'autre part, l'étude par diagrammes d'états a mis en évidence une phase d'équilibrage caractéristique. Il serait alors possible d'accélérer grandement celle-ci par le biais d'une prédiction du point d'équilibre. On pourrait ainsi rajouter des informations de feed-back au niveau protocolaire pour ainsi calculer une estimation de la prédiction du point d'équilibre.

ANNEXE

CODE SCILAB DE LA MODÉLISATION TDCN

Listing 1: Modelisation Scilab

```
function [CLA2]=tdcn(CLA,LLA,A,B,BW)
          if sum(CLA)>=BW then
                    CLA2=CLA-A.*(CLA./LLA)
          else
                    CLA2=CLA+B.*(LLA./CLA)
          end
endfunction

function [POINT]=Calc_Convergence(LLA,D)
          POINT=LLA/sum(LLA).*D;
endfunction

function [DIST]=distance(MATRICE,POINT)
          Size=size(MATRICE);
          MPTS=ones(Size(1),1)*POINT;
          CARRE=(MATRICE-MPTS).^2;
          DIST=sqrt(sum(CARRE,'c'));
endfunction

function create_Trace(FILE_TRACE,M,N)

         unix("rm "+ FILE_TRACE)
         u=file('open',FILE_TRACE);
          if M==0 then
                    write(u,CLA);
          end
          for i =1:N,
                    CLA=tdcn(CLA,LLA,A,B,BW);
                    if i>=M then
                              write(u,CLA);
                    end
          end
          file('close',u);
endfunction
```

Listing 2: Fonction d'affichage graphique

```
function Graph_Dist(MATRICE,DEBUT,POINT,COLOR)
          xtitle("Distances au point de convergence a partir de t="+
              string(DEBUT),"temps en seconde","Distance(en Kb/s)");
          SIZE=size(MATRICE);
          DISTANCES=distance(MATRICE(DEBUT:SIZE(1),:),POINT);
          X=size(DISTANCES);
          X=1:1:X(1);
          plot2d (X,DISTANCES,style=COLOR);
endfunction

function Graph_Dist_log(MATRICE,DEBUT,POINT,COLOR)
        SIZE=size(MATRICE)
          DISTANCES=distance(MATRICE(DEBUT:SIZE(1),:),POINT);
          X=size(DISTANCES);
          X=1:1:X(1);
          plot2d(X,DISTANCES,style=COLOR,logflag="nl");
endfunction

function [R]=fenetre(M,F);
          Size=size(M);
          for i =1:Size(1),
                    if i<F then
                            R(i)=sum(M(1:i))/i;
                    else
                            R(i)=sum(M((i-F+1):i))/F;
                    end
          end
endfunction

function Graph_Dist_Moy_log(MATRICE,DEBUT,POINT,FEN,COLOR)
        SIZE=size(MATRICE);
        DISTANCES=distance(MATRICE(DEBUT:SIZE(1),:),POINT);
          DIST_F=fenetre(DISTANCES,FEN);
          X=size(DISTANCES);
          X=1:1:X(1);
          plot2d(X,DIST_F,style=COLOR,logflag="nl");
endfunction

function [REP]= normal_inv(DER,SUM,DEG)
          MOY=SUM/DEG;
          REP=grand(1,DEG-1,'nor',MOY,DER);
          REP=[REP SUM-sum(REP)];
endfunction
```

Listing 3: Fonction sur les diagrammes de phase

```
function Graph_phase(MATRICE,DEBUT,COLOR)
          xtitle("Diagramme de phase a partir de t="+string(DEBUT),"CLA2
              (en Kb/s)","CLA1(en Kb/s)");
          plot2d(MATRICE(:,1),MATRICE(:,2),style=COLOR);
endfunction

function [R]=Espace_phase_G2(MATRICE,P,taille)
          D=distance(MATRICE,P);
          DF=fenetre(D,taille);
          ERR=min(DF);
          R=1.1*ERR;
          i=1;
          while DF(i)>R do i=i+1; end
          TPS=i;
          R =[ERR TPS];
endfunction

function [R]=Espace_phase_G(MATRICE,P)
        D=distance(MATRICE,P);
          ERR=min(D);
          i=1;
          while D(i)>ERR do i=i+1; end
          TPS=i;
          R =[ERR TPS];
endfunction
```

Listing 4: Diagramme de distance simple

```
CLA=[100000 100000 100000];
LLA=[30000 20000 10000];
A=[100 100 100];
B=[100 100 100];
BW=80000;
xbasc();
xtitle("Diagramme de Distance","Temps en sec","Distance en kb/s");
k=10;
n=1;
color=2;
WIN=20;
create_Trace('theorique.dat',n,2500);
M=read('theorique.dat',-1,3);
P=Calc_Convergence(LLA,BW);
Graph_Dist_log(M,n,P,color);
Graph_Dist_Moy_log(M,n,P,WIN,color+1);
```

Listing 5: Etat du TDCN en fonction de A et B

```
CLA=[100000 100000 100000];
LLA=[10000 20000 30000];
BW=80000;
 xbasc();
xtitle("Diagramme","Erreur de convergence en Kb/s","Delai de convergence
    en seconde");
n=2;
Duree=2000;
for i = list(100,500,1000,2000),
        m=1;
        for j=i/20:i/10:5*i,
                A=ones(1,3)*j;
                B=ones(1,3)*i;
                create_Trace('theorique.dat',n,Duree);
                M=read('theorique.dat',-1,3);
                P=Calc_Convergence(LLA,BW);
                D=Espace_phase_G2(M,P,100);
                if D(2)< (Duree-200) then
                        if D(1)<4000 then
                                if m>2 then
                                        C=[C;D];
                                else
                                        C=D;
                                        m=m+1;
                                end
                        end
                end
        end
        plot2d(C(:,1),C(:,2),style=n)
        n=n+1;
        clear C;
end
```

Listing 6: Etat du TDCN en fonction du degré

```
D0=100000;
C0=100000;
L0=500000;
A0=10000;
B0=10000;
BW=1000000;
xbasc();
 n=1;
m=1;
h=1;
Duree=1100;
for i = 10:50:610,
        m=m+1;
        CLA=ones(1,i)*D0/sqrt(i)+BW/i;
        LLA=ones(1,i)*L0/i;
        A=ones(1,i)*A0/i;
        B=ones(1,i)*B0/i;
        create_Trace('theorique.dat',n,Duree);
        M=read('theorique.dat',-1,i);
        P=Calc_Convergence(LLA,BW);
        D=Espace_phase_G2(M,P,300);
        if D(1)<4000 then
                if h>1 then
                        C=[C;i,D];
                else
                         C=[i,D];
                        h=h+1;
                end
        end
end
xtitle("Erreur de convergence en fonction du degre ","degre du systeme",
    "Delai de convergence en seconde");
plot2d(C(:,1),C(:,2),style=3);
```

Listing 7: Etat du TDCN en fonction de la dispersion des CLA

```
degre=50;
BW=1000000;
somhla=2000000;
somlla=500000;
LLA=ones(1,degre)*somlla/degre;
m=1;
h=2;
xbasc();
  xtitle("Delai en fonction de l ecart-type","Ecart-type des CLAs en Kb/
      s","Delai de convergenceen seconde");
for  ecart =2000:1000:15000,
        m=m+1;
        clear D;
        nb=1;
        for rep =1:1:20,
                CLA=normal_inv(ecart,somhla-somlla,degre)+LLA;
                A=ones(1,degre)*somlla/degre/100;
                B=ones(1,degre)*somlla/degre/100;
                create_Trace('theorique.dat',1,3000);
                M=read('theorique.dat',-1,degre);
                P=Calc_Convergence(LLA,BW);
                if isdef('D') then
                         E=Espace_phase_G2(M,P,100);
                        if E(1)<10000 then
                                D=(D.*nb+E)./(nb+1);
                                nb=nb+1;
                        end
                else
                        E=Espace_phase_G2(M,P,100);
                        if E(1)<10000 then
                        D=E;
                end
        end
end
if h>2 then
        C=[C;ecart,D];
else
        C=[ecart,D];
        h=h+1;
end
 plot2d(C(:,1),C(:,3),style=3);
```

PUBLICATIONS

Les travaux réalisés pendant ma thèse ont donné lieu aux publications suivantes :

OLIVIER FERVEUR. *procédé de gestion d'un réseau de transmission de type mpls*. BREVET FRANÇAIS DÉPOSÉ À L'INPI LE 28 AVRIL 2008 SOUS LE N°08 02392.

OLIVIER FERVEUR, E. GNAEDINGER, F. LEPAGE AND R. KOPP. *proposition d'une nouvelle régulation d'accès au coeur de réseau des opérateurs d'infrastructure de communication*. 5ÈME CONFÉRENCE INTERNATIONALE FRANCOPHONE D'AUTOMATIQUE, CIFA'2008, ROUMANIE (2008)

OLIVIER FERVEUR, E. GNAEDINGER, F. LEPAGE AND R. KOPP. *new sharing access regulation for core's network*. 11TH IEEE SINGAPORE INTERNATIONAL CONFERENCE ON COMMUNICATION SYSTEMS, ICCS 2008, GUANGZHOU (2008).

OLIVIER FERVEUR, E. GNAEDINGER, F. LEPAGE AND R. KOPP. *new quality of service policy for wholesale approach*. IN PROCESS. IEEE/ACM TRANSACTIONS ON NETWORKING.

BIBLIOGRAPHIE

[1] *Network Working Group Praveen Muley Internet Draft Mustapha Aissaoui Expires : August 2008 Matthew Bocci Pranjal Kumar Dutta*, 2009.

[2] *Network Working Group Y. Kamite, Ed. Internet-Draft Y. Wada Expires : September 7, 2006 NTT Communications Y. Serbest AT&T*, 2009.

[3] L. Andersson and T. Madsen. Provider Provisioned Virtual Private Network (VPN) Terminology. RFC 4026 (Informational), March 2005.

[4] Guido Appenzeller, Isaac Keslassy, and Nick McKeown. Sizing router buffers. *SIGCOMM Comput. Commun. Rev.*, 34(4) :281–292, 2004.

[5] Ron Banner and Ariel Orda. Multipath routing algorithms for congestion minimization. *IEEE/ACM Trans. Netw.*, 15(2) :413–424, 2007.

[6] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss. An Architecture for Differentiated Service. RFC 2475 (Informational), December 1998. Updated by RFC 3260.

[7] M. Bocci, I. Cowburn, and J. Guillet. Network high availability for ethernet services using IP/MPLS networks. *Communications Magazine, IEEE*, 46(3) :90–96, 2008.

[8] A. Bosco, R. Mameli, E. Manconi, and F. Ubaldi. Edge distributed admission control in MPLS networks. *Communications Letters, IEEE*, 7(2) :88–90, 2003.

[9] B. Braden, D. Clark, J. Crowcroft, B. Davie, S. Deering, D. Estrin, S. Floyd, V. Jacobson, G. Minshall, C. Partridge, L. Peterson, K. Ramakrishnan, S. Shenker, J. Wroclawski, and L. Zhang. Recommendations on Queue Management and Congestion Avoidance in the Internet. RFC 2309 (Informational), April 1998.

[10] B. Briscoe. Flow rate fairness : dismantling a religion. *ACM SIGCOMM Computer Communication Review*, 37(2) :63–74, 2007.

[11] K. Chan, R. Sahita, S. Hahn, and K. McCloghrie. Differentiated Services Quality of Service Policy Information Base. RFC 3317 (Informational), March 2003.

[12] C.N. Chuah. *A Scalable Framework for IP-network Resource Provisioning Through Aggregation and Hierarchical Control*. Computer Science Division, University of California, 2001.

[13] Clark D. and Wroclawski J. *An Approach to Service Allocation in the Internet*. IETF Draft Report, Massachusetts Institute of Technology, July 1997.

[14] J. de Rezende. Assured service evaluation. In *IEEE Globecomm'99, Rio de Janeiro, Brazil*, December 1999.

[15] AF Filippov. *Differential equations with discontinuous righthand sides*. Kluwer Academic Pub, 1988.

[16] C. Fraleigh, F. Tobagi, and C. Diot. Provisioning IP backbone networks to support latency sensitive traffic. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, volume 1, 2003.

[17] B. Gleeson, A. Lin, J. Heinanen, G. Armitage, and A. Malis. A Framework for IP Based Virtual Private Networks. RFC 2764 (Informational), February 2000.

[18] J. Heinanen and R. Guerin. A Two Rate Three Color Marker. RFC 2698 (Informational), September 1999.

[19] T. Ikeda, S. Sampei, and N. Morinaga. TDMA-based adaptive modulation with dynamic channel assignment forhigh-capacity communication systems. *Vehicular Technology, IEEE Transactions on*, 49(2) :404–412, 2000.

[20] Dina Katabi, Mark Handley, and Charlie Rohrs. Internet congestion control for future high bandwidth-delay product environments. In *SIGCOMM '02 : Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 89–102, New York, NY, USA, 2002. ACM.

[21] M. Kodialam and TV Lakshman. Dynamic routing of restorable bandwidth-guaranteed tunnels using aggregated network resource usage information. *IEEE/ACM Transactions on Networking (TON)*, 11(3) :399–410, 2003.

[22] S. Köhler and A. Binzenhöfer. MPLS traffic engineering in OSPF networks-a combined approach. In *Proceedings of ITC*, volume 18, 2003.

[23] K. Kompella and J. Lang. Procedures for Modifying the Resource reSerVation Protocol (RSVP). RFC 3936 (Best Current Practice), October 2004.

[24] M. Lasserre and V. Kompella. Virtual Private LAN Service (VPLS) Using Label Distribution Protocol (LDP) Signaling. RFC 4762 (Proposed Standard), January 2007.

[25] Wei Lin, Rong Zheng, and Jennifer C. Hou. How to make assured service more assured. In *ICNP '99 : Proceedings of the Seventh Annual International Conference on Network Protocols*, page 182, Washington, DC, USA, 1999. IEEE Computer Society.

[26] Liu Zhen Malouch Naceur. Rr-4469 - performance evaluation of rio routers. Technical report, INRIA, Juin 2002.

[27] Zoubir Mammeri. Framework for parameter mapping to provide end-to-end qos guarantees in intserv/diffserv architectures. *Computer Communications*, 28(9) :1074 – 1092, 2005.

[28] P. Marbach. Priority service and max-min fairness. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 1, 2002.

[29] A. Markopoulou, G. Iannaccone, S. Bhattacharyya, C. Chuah, and C Diot. Characterization of failures in an ip backbone. *IEEE INFOCOM*, 2004.

[30] Petter Mosebekk. *A linux implementation and analysis of the eXplicit Control Protocol (XCP).* PhD thesis, University of Oslo, May 2005.

[31] K. Papagiannaki, R. Cruz, and C. Diot. Network performance monitoring at small time scales. In *Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement*, pages 295–300. ACM New York, NY, USA, 2003.

[32] P. Pongpaibool. Characteristics of Internet Traffic in Thailand, 2006.

[33] K. Ramakrishnan, S. Floyd, and D. Black. The Addition of Explicit Congestion Notification (ECN) to IP. RFC 3168 (Proposed Standard), September 2001.

[34] S. Sarpturk, Y. Istefanopulos, and O. Kaynak. On the stability of discrete-time sliding mode control systems. *IEEE Transactions on Automatic Control*, 32(10) :930–932, 1987.

[35] N. Seddigh, B. Nandy, and P. Pieda. *Bandwidth Assurance Issues for TCP Flows in a Differentiated Services Network*. IEEE Globecomm'99, Rio de Janeiro, Brazil, 1999.

[36] P. Sousa, M. Rocha, M. Rio, and P. Cortez. Automatic provisioning of QoS aware OSPF configurations, 2007.

[37] P. Tinnakornsrisuphap and AM Makowski. Limit behavior of ECN/RED gateways under a large number of TCP flows. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications Societies. IEEE*, volume 2.

[38] P. Trimintzios, I. Andrikopoulos, G. Pavlou, C. Cavalcanti, D. Goderis, Y. T'Joens, P. Georgatsos, L. Georgiadis, D. Griffin, C. Jacquenet, et al. An Architectural Framework for Providing QoS in IP Differentiated Services Networks. In *7th IFIP/IEEE International Symposium on Integrated Network Management (IM 2001)*, 2001.

[39] Christo Wilson, Chris Coakley, and Ben Y. Zhao. Fairness attacks in the explicit control protocol. *Quality of Service, 2007 Fifteenth IEEE International Workshop on*, pages 21–28, June 2007.

[40] O. Younis and S. Fahmy. Constraint-Based Routing in the Internet : Basic Principles and Recent Research. *IEEE Communications Surveys and Tutorials*, 5(1) :2–13, 2003.