



THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Informatique

École doctorale Matisse

présentée par

Selim ELLOUZE

préparée à l'unité de recherche et développement
Orange Labs Lannion

**Mécanismes de
collaboration entre
réseaux et
services
applicatifs pour
l'optimisation des
ressources et des
services**

**Thèse à soutenir à Paris
le 2 Juillet 2013**

devant le jury composé de :

Prof. Damien MAGONI (RAPPORTEUR)

Professeur à l'Université de Bordeaux 1

Prof. Houssam AFIFI (RAPPORTEUR)

Professeur à Télécom SudParis

Dr. Houda LABIOD (EXAMINATRICE)

Enseignante Chercheur à Télécom Paristech

Dr. Fabrice GUILLEMIN (EXAMINATEUR)

Chercheur senior à Orange Labs

Prof. Omar CHERKAOUI (EXAMINATEUR)

Professeur à l'Université de Québec

Prof. Gerardo RUBINO (DIRECTEUR DE THÈSE)

Professeur à l'Université de Rennes 1

Dr. Tayeb LEMLOUMA (CO-DIRECTEUR DE THÈSE)

Enseignant Chercheur IRISA

À mes parents,

Remerciements

J'exprime ma gratitude à Tayeb Lemlouma et Gerardo Rubino pour m'avoir accordé leur confiance en acceptant l'encadrement et la direction de cette thèse. Je remercie également Jean-François Jestin de m'avoir encadré et soutenu durant cette période à Orange Labs. Leurs conseils, leurs implications et leurs disponibilités sont exceptionnels.

Je présente également mes remerciements aux membres du jury, Houda Labiod, Fabrice Guillemin et Omar Cherkaoui qui ont accepté d'évaluer mon travail. Je remercie vivement Damien Magoni et Houssam Afifi qui m'ont accordé l'honneur d'être rapporteurs de ma thèse.

Je suis aussi très reconnaissant à mes anciens collègues Gaël Fromentoux, Emile Stephan, Bertrand Mathieu et Joel Penhoat pour leur aide précieuse, leur soutien continu, leur humour et leurs qualités humaines et professionnelles. Ces années de thèse se sont passées très vite en leur présence et l'expérience que j'ai acquise à leur côté est un témoin de leur valeur. Mes remerciements vont aussi à tous les membres de mon ancienne unité, Xavier Marjou, Pascal Rousset, Luc Le Beller, Nathalie Omnes, Frédéric Fieau, Nicolas Bihannic et à ceux de l'unité SINA. Le petit café du matin est un rituel de bonne humeur et de convivialité que ces personnes entretiennent avec normalité.

Je tiens à remercier aussi tous mes proches et mes amis, Angela qui m'a accompagné un bout de chemin, Rosa et Oscar qui m'ont accordé un soutien sans faille aux moments difficiles, Bettina, Ambroise, Josselin, Adrian, Bayrem, Aymen pour leur présence, leur bonne humeur et leurs encouragements.

Mes dernières pensées iront à ma famille. Mes frères et soeurs, Hanen, Sonia, Sofien et Mahmoud, tous docteurs, ne m'ont guère laissé le choix mais étaient toujours présents. Je suis extrêmement reconnaissant à Jennifer qui a vécu cette thèse avec moi aux moments difficiles avant les moments de joie. Je remercie énormément mon neveu Anes et mon beau frère Larvi pour leur présence à tout moment à mes côtés. Ma plus profonde gratitude va à mes parents pour leur foi et leurs sacrifices. Cette thèse est l'accomplissement d'un chemin dont les origines naissent dans la grandeur des valeurs qu'ils m'ont transmises.

Enfin, je remercie tous ceux, qui de près ou de loin, ont contribué à la réalisation de ce travail.

Résumé

Dans cette nouvelle ère du numérique, l'accès à l'information est entré désormais dans une autre dimension. Nous assistons à la dominance d'un modèle fondé sur les opportunités offertes par un accès mondialisé à l'Internet et à son application phare : le "*World Wide Web*". Les services se sont multipliés. Les terminaux se sont diversifiés. Les technologies de transport se sont améliorées. Les attentes se sont élevées. Dans cette spirale que nous nous abstenons de qualifier, les opérateurs se trouvent désormais confrontés à une croissance soutenue du trafic dans leurs réseaux, en grande partie due au transport de flux vidéo. Les fournisseurs de services sur Internet se trouvent aussi concernés par la problématique de la qualité de service dont dépend la satisfaction de leurs utilisateurs. Pour l'ensemble des acteurs, ces nouvelles tendances présentent à la fois des défis et des opportunités. Les défis se concentrent dans la problématique de gestion de la demande croissante du trafic tout en maintenant une qualité d'expérience appropriée pour les utilisateurs. Les opportunités proviendront de l'adéquation entre une demande croissante des services Web en termes de qualité de services et des ressources qui devront supporter la distribution de ces services. Il est crucial pour chaque acteur de bien se positionner dans la chaîne de valeur pour gérer cette adéquation. Le rôle que prendra le réseau support, simple ensemble de tuyaux surdimensionnés, ou bien réseau intelligent offrant des fonctions avancées de contrôle illustre parfaitement cet enjeu. Ces deux alternatives sont respectivement connues sous les termes "*dumb-pipe*" ou "*smart network*".

Dans cette thèse, nous considérons une nouvelle approche, qui se veut simple, efficace et adaptée pour faire face à ces défis. Les opérateurs réseaux et les fournisseurs de services sont mutuellement gagnants dans l'amélioration du transport de données dans les réseaux tout en continuant à opérer leur propre infrastructure. Cette démarche coopérative est le point de départ de nos travaux qui visent à définir un cadre, une architecture et des techniques appropriées qui amèneront ces acteurs à collaborer en vue de gérer conjointement cette problématique. Cette collaboration est nécessaire car chaque acteur quoique prisonnier de ses contraintes peut les transformer en relations contractuelles dans un processus client fournisseur pour l'optimisation de la gestion du trafic.

Notre effort a d'abord porté sur la définition d'un cadre de collaboration adapté. La collaboration reposera sur une communication efficace entre les acteurs pour partager leurs contraintes et exposer leurs besoins. Nous avons développé à cet égard une interface qui offre d'intéressantes facilités d'intégration et d'évolution tout en masquant la complexité de l'éco-système propre à chaque acteur et en préservant leur savoir-faire et leurs atouts respectifs. Cette interface est le pivot d'une architecture fonctionnelle conçue pour résoudre cette problématique dans une démarche coopérative. Des entités fonctionnelles précises, distribuées entre la couche réseau et la couche service composent cette architecture. Le processus d'optimisation inter-couches qui repose sur l'interface

de collaboration est implémentée via une architecture technique issue de la solution fonctionnelle.

Notre contribution s'est focalisée par la suite sur les techniques d'optimisation. Nous avons étudiée en premier lieu une approche de collaboration que nous appelons passive car les opérateurs réseaux ne prennent pas part au processus décisionnel. Nous nous sommes intéressés à la pertinence des informations que peuvent fournir ces derniers pour guider les fournisseurs de services dans l'optimisation de l'orientation du trafic. Les décisions prises au niveau de la couche applicative définissent les connections à établir entre les entités applicatives en fonction des conditions des chemins réseaux qui les relient. En second lieu, nous avons considéré l'adoption d'une approche de collaboration active qui permet aux opérateurs réseaux de prendre part activement au processus d'optimisation partant du fait que les opérateurs sont capables de fédérer la gestion des besoins de tous les services applicatifs qui utilisent leurs réseaux. Cette approche repose sur un échange bilatéral d'informations entre les acteurs service et réseau. L'optimisation est déléguée aux opérateurs réseaux qui prennent en compte leurs contraintes mais aussi les besoins des différents services applicatifs qui collaborent pour fournir une réponse personnalisée et optimale à chacun dans le but de réaliser une gestion conjointe de leurs trafics.

Pour valider les différentes approches et techniques que nous avons proposées, nous avons réalisé des simulations et des expériences dans des réseaux de test pour évaluer les performances des métriques réseaux et services. Nous avons développé des modèles analytiques et nous les avons implémentés avec le logiciel Matlab pour réaliser les simulations. Nous avons développé aussi un prototype de distribution de contenus pour tester le processus en environnement réel. Les résultats obtenus ont démontré que nos propositions permettent de réaliser une meilleure répartition de la charge dans le réseau tout en diminuant le taux d'utilisation global du réseau. Une amélioration de la qualité de service est assurée par une meilleure disponibilité des ressources dans le réseau. Ces résultats confirment l'intérêt d'introduire les approches de collaboration dans le paysage actuel pour apporter une réponse efficace à la problématique de croissance du trafic, des contraintes de qualité des services et d'exigences des utilisateurs

Abstract

In this new digital world, driven by the dominance of a model based on the opportunities offered by global access to the Internet and its killer app : the World Wide Web, access to information is becoming a matter of a good experience and responsiveness. We are witnessing the Web services are of increasing popularity. New types of terminals are proposed. Communications technologies are improved. Users' expectations are becoming higher. In such a context, network operators are facing serious challenges arising from the management of a massive traffic growth, largely driven by the increasing amount of video streams while internet services providers are also concerned by the issue of providing an adequate quality of experience to their end-users. For both actors, these dual trends present both challenges and opportunities. The challenges arise from the issues of managing the growing demand for traffic while maintaining appropriate quality of experience for users. Opportunities come from a smart management of the increasing demands of Web services in terms of quality of services and of the resources that will support the delivery of these services. It is then crucial for each actor to be well-positioned in the value chain to take part in this process. The role that will be played by the networks, as a basic set of oversized pipes, or as an intelligent network providing advanced management facilities, perfectly illustrates this issue. These two alternatives are respectively known as the "dumb-pipe" or "smart networks".

In this thesis, we consider a new approach, which is simple, effective and adapted to meet these challenges. Network operators and service providers can mutually benefit from improving the data delivery in the networks while continuing to fully control their infrastructures. This collaborative approach is the starting bloc of our work aiming at defining a framework, an architecture and appropriate procedures to bring these actors to work together to manage this problem. This collaboration is particularly necessary because each actor, though prisoner of its constraints and capacities, can turn them into a contractual relation with the other in a client-supplier process for the optimization of traffic management.

Our efforts were first devoted to the definition of a suitable framework for the collaboration. We promoted a collaboration process based on effective exchange of information between stakeholders to share their knowledge and expose their constraints. We developed an interface to enable such a process while not only offering interesting facilities for integration and evolution but also hiding the complexity of the eco-systems of the actors and protecting their know-how and strengths. This interface is the backbone of a functional architecture designed to address the above mentioned issue in a collaborative approach. Functional entities composing this architecture are defined within the network layer and the service layer to play specific roles. The process of cross-layer optimization based on the collaboration interface is implemented via a technical architecture after the functional solution.

Our contribution was devoted also to the specification of efficient optimization procedures. We first studied a collaborative approach, we call passive as network operators do not take part in the decision process for the overlay traffic management. We investigated the relevance of different types of metrics they can provide to guide service providers in optimizing the management of their traffic. Decisions made at the overlay layer define the connections established between overlay entities depending on the conditions of network paths between them. In a different approach, we considered an active collaboration which allows networks operators to take part actively in the optimization process. Indeed, network operators are able to unite the processing of the information related to different overlay services that use their networks. This approach relies on a mutual exchange of information between network operators and internet service providers. The optimization is delegated to the network layer which considers the capacities and the constraints of all collaborating actors. It provides them then with a personalized and optimal response to address the traffic management issue in order to achieve a mutual a common benefit.

in order to evaluate the different approaches and techniques that we propose, we carried different simulations and experiments in real networks to measure the performances of the network and service layers. We have developed analytical models that we implemented within the Matlab software to perform the simulations. We have also developed a prototype of a content distribution service to evaluate our optimization framework in a real network. The results showed that our proposals can achieve better load balancing in the network while minimizing the overall utilization of the network. Improving the quality of service is ensured by improved the availability of network resources. These encouraging results show the benefits of introducing collaborative approaches to efficiently address the traffic growth issues and quality of experience expectations.

Publications

Collaboration between Networks and Applications in the Future Internet
Selim Ellouze, Bertrand Mathieu, Toufik Ahmed et Nico Schwan Media Networks :
Architectures, Applications, and Standards, Chapitre 7 Auerbach Publications, 2011.

Improving End-to-End QoE via Close Cooperation between Applications and ISPs Bertrand Mathieu, Selim Ellouze, Nico Schwann, Eleni Myonkyati IEEE Communications Magazine Mars 2011.

A proposal for improving network-aided CDN server Selim Ellouze, Bertrand Mathieu, Tayeb Lemlouma CFIP/NOTERE, 2012.

Collaboration Interface for Network Services Invocation by Cloud Management Systems Bertrand Mathieu, Abdelhadi Chari, Selim Ellouze, Irène Grosclaude IEEE CloudNet 2012.

A Bidirectional Network Collaboration Interface for CDNs and Clouds Services Traffic Optimization Selim Ellouze, Bertrand Mathieu, Tayeb Lemlouma IEEE International Conference on Communications, 2013.

Brevets

PROCEDE DE COMMUNICATION POUR LA DISTRIBUTION HYBRIDE DE DONNEES Selim Ellouze, Gael Fromentoux, Emile Stephan. Numéro : FR20100058857 20101027/WO 2012056176 (A2). Date de publication : 04-05-2012.

TRANSMISSION DE FLUX DE DONNEES ADAPTABLE Selim Ellouze, Gael Fromentoux, Xavier Marjou. Numéro : FR20100052403 20100331 ; WO2011FR50711 20110330. Date de publication : 06-02-2013.

TECHNIQUE DE DISSIMULATION AVEC DISTINCTION DES SOURCES DE DONNEES ET OPTIMISATION DE LEUR TRANSPORT Selim Ellouze, Gael Fromentoux, Emile Stephan. Numéro : En cours d'attribution.

EN COURS DE TRAITEMENT Selim Ellouze, Emile Stephan.

Normalisation

ALTO session for CDN Interconnection draft-stephan-cdni-alto-session-ext-03 Emile Stephan, Selim Ellouze. Groupe de travail IETF CDN-i. Date de publication : 16-04-2013.

Table des matières

Table des matières	3
Table des figures	9
Liste des tables	13
Introduction	17
0.1 Contexte et problématique	17
0.2 Motivations, approche, et défis	18
0.3 Contributions	19
0.4 Organisation du rapport	20
1 Analyse des différentes interfaces de communications entre services et réseaux pour la gestion des flux de données	23
1.1 ALTO, l'interface de référence	24
1.1.1 Le projet P4P	24
1.1.1.1 Le système P4P	25
1.1.1.2 Le plan de commande du système P4P	25
1.1.2 Le protocole ALTO	27
1.1.2.1 La carte du réseau	27
1.1.2.2 La matrice des coûts	29
1.2 Les interfaces proposées dans la littérature	30
1.2.1 Des informations réseau pour une meilleure gestion du trafic au niveau applicatif	30
1.2.1.1 PaDIS	31
1.2.1.2 CaTE	31
1.2.1.3 DONAR	32
1.2.1.4 Le projet NAPA-WINE	34
1.2.1.5 I-CDN	36
1.2.1.6 Les systèmes de coordonnées virtuelles	36
1.2.2 La modulation du Traffic Engineering aux contraintes des applications	38
1.2.3 Synthèse des techniques étudiées	39
1.3 Les interfaces issues du monde de la normalisation	39

1.3.1	Le Next Generation Network	39
1.3.1.1	Architecture fonctionnelle du NGN	41
1.3.1.2	Le contrôle du transport dans le NGN	41
1.3.2	L'IP Multimedia Subsystem	45
1.3.2.1	Présentation de l'IMS	45
1.3.2.2	L'architecture 3GPP PCC	45
1.3.3	L'architecture commune 3GPP et TISPAN pour l'intégration de l'IMS et le NGN	47
1.3.4	OSA/Parlay	48
1.3.4.1	Introduction à OSA/Parlay	48
1.3.4.2	Présentation des APIs	48
1.3.4.3	OSA/Parlay Connectivity Manager	49
1.3.5	GSMA OneAPI	50
1.3.5.1	Introduction à OneAPI	50
1.3.5.2	Les APIs OneAPI	50
1.3.6	Openflow	51
1.4	L'utilité d'une interface de collaboration	52
1.4.1	Des interfaces inappropriées	52
1.4.1.1	Les limites de la maîtrise de la QoS dans le réseau	52
1.4.1.2	Un intérêt convergent vers la collaboration entre réseaux et services dans la littérature	53
1.4.2	L'utilité d'un système de collaboration entre réseaux et services	54
1.4.2.1	L'intérêt pour les services CDNs	54
1.4.2.2	L'intérêt pour les applications P2P	57
1.4.2.3	Impacts stratégiques et financiers pour les opérateurs réseaux	58
1.5	Conclusion	59
2	L'approche collaborative entre réseaux et services	61
2.1	L'interface de collaboration entre réseaux et applications	61
2.1.1	Les exigences de l'approche collaborative	62
2.1.2	L'architecture du système de collaboration	63
2.1.2.1	Présentation de l'architecture	63
2.1.2.2	Présentation des modules fonctionnels	63
2.1.3	L'interface CINA : Collaboration Interface Between Networks and Application	67
2.1.3.1	Présentation	67
2.2	L'approche de collaboration passive des opérateurs réseaux	67
2.2.1	Introduction de l'approche	67
2.2.2	Application de l'approche	68
2.2.2.1	Le contexte réseau pour les services CDNs	68
2.2.3	ALTO pour interface	69
2.2.4	Les différentes classes d'information pour l'interface ALTO	70

2.2.4.1	Les défis inhérents à la collecte et exposition des informations	70
2.2.4.2	La classe d'informations distance	72
2.2.4.3	La classe d'informations latence	74
2.2.4.4	La classe d'informations bande passante	74
2.2.4.5	Autres classes d'informations	75
2.2.5	Proposition d'une nouvelle classe d'informations	76
2.2.5.1	La classe d'information coût de transport	76
2.2.5.2	La fonction de coût	79
2.2.5.3	Analyse des informations exposées par les coûts de transport	80
2.3	L'approche de collaboration active des opérateurs réseaux	83
2.3.1	Limites de l'approche passive	83
2.3.2	Proposition d'amélioration de l'approche passive	85
2.3.2.1	La matrice de proportions	85
2.3.2.2	La matrice de contraintes	86
2.3.3	L'échange bidirectionnel d'informations	87
2.3.4	Le problème d'optimisation	88
2.3.4.1	Modélisation mathématique du problème	89
2.3.4.2	Définition du problème	91
2.3.4.3	La résolution du problème	93
3	Application des approches de collaboration	95
3.1	Application de l'approche aux services CDN	95
3.1.1	Présentation d'une plateforme CDN	95
3.1.1.1	La plateforme d'Akamai	96
3.1.2	Intégration des modules de collaboration dans une plateforme CDN	98
3.2	Application de l'approche aux applications P2P	100
3.2.1	Intégration dans les architectures P2P existantes	100
3.2.2	Proposition d'une architecture P2P dédiée	101
3.2.2.1	Intérêt de cette conception	101
3.2.2.2	Conception du système	102
3.2.2.3	Structure des données	103
3.2.2.4	Structure de l'overlay	103
3.2.2.5	L'approche retenue pour les échanges de données	104
3.2.3	L'optimisation du rendement du système	105
3.2.3.1	Notations et conditions	105
3.2.3.2	Le modèle d'échange de données	105
3.2.3.3	L'optimisation du rendement du système	106
3.2.4	Les échanges de données dans un groupe de pairs	107
3.2.4.1	Le schéma de routage pour un groupe complet	108
3.2.5	La formation d'un overlay de groupes P2P	111
3.2.5.1	Exploitation des ressources additionnelles	111
3.2.5.2	Atténuation des instabilités de groupes dans l'overlay	114

3.2.6	La construction de l'overlay P2P	116
3.2.6.1	Connexions intra-groupe	116
3.2.6.2	Connexions entre groupes dans le même PID	116
3.2.6.3	Connexions entre groupes dans des PIDs différents	117
4	Evaluations des performanes des approches de collaboration	119
4.1	Les simulations analytiques	119
4.1.1	Les principaux défis relatifs aux évaluations	119
4.1.2	Introduction au modèle	120
4.1.2.1	Notations	121
4.1.3	Présentation du modèle	122
4.1.3.1	Les conditions du réseau BCMP	123
4.1.3.2	Les paramètres du modèle analytique	123
4.1.3.3	La stabilité du système	123
4.1.3.4	Les mesures à l'état d'équilibre	125
4.1.3.5	Les métriques de performance	125
4.1.4	La configuration du modèle analytique pour les simulations	126
4.1.4.1	Implémentation du modèle	126
4.2	Évaluation de l'approche passive de collaboration	127
4.2.1	Présentation des paramètres	127
4.2.2	Résultats des simulations	129
4.2.3	Synthèse et conclusions	134
4.3	Évaluation de l'approche active de collaboration	137
4.3.1	Présentation des paramètres	137
4.3.2	Résultats des simulations	138
4.3.3	Synthèse et conclusions	143
4.4	Évaluation de l'application de l'approche de collaboration à un système P2P	145
4.4.1	Présentation des paramètres de la simulation	145
4.4.2	Résultats des simulations	146
4.4.3	Synthèse et conclusions	151
4.5	Évaluation de l'approche collaborative sur un prototype d'un système de distribution de contenus	154
4.5.1	Présentation du prototype de test	154
4.5.1.1	Le prototype CDN	154
4.5.1.2	Le prototype CINA	154
4.5.1.3	Le module de monitoring réseau	158
4.5.2	Le réseau de test	158
4.5.3	Présentation de la plateforme de test	158
4.5.3.1	La configuration des services CDN	159
4.5.3.2	Les résultats des redirections des clients	159
4.5.4	Les résultats des tests	162
4.5.4.1	La mesure des débits applicatifs de chaque PID	162
4.5.4.2	La mesure d'autres indicateurs de performance	165

4.5.4.3	Analyse de la répartition de la charge dans le réseau . . .	168
4.5.5	Conclusion et synthèse	169
Conclusion		171
Bibliographie		183

Table des figures

1.1	Illustration de l'architecture du système de contrôle P4P	26
1.2	Illustration de l'architecture fonctionnelle d'ALTO	27
1.3	Illustration des services ALTO disponibles à travers le protocole ALTO .	28
1.4	Illustration d'une carte réseau	28
1.5	Illustration d'une carte réseau	29
1.6	Illustration d'une matrice de coûts	30
1.7	Illustration de l'intégration du système PaDIS	31
1.8	Illustration du modèle système DONAR	33
1.9	Illustration de l'architecture globale du projet NAPA-WINE	35
1.10	Illustration des interactions entre les systèmes TE et SS	39
1.11	L'architecture fonctionnelle du système NGN	42
1.12	Illustration des interfaces externes du NASS	43
1.13	Illustration des interfaces externes du RACS	44
1.14	Illustration de l'architecture commune 3GPP/TISPAN	47
1.15	Illustration des interfaces prodiguées par les APIs OMA entre services et réseaux	50
1.16	Illustration de la diversité de choix pour la sélection d'un noeud de distribution	56
2.1	L'architecture fonctionnelle du système de collaboration entre réseaux et applications	64
2.2	La diversité de chemins dans le contexte réseau	69
2.3	Illustration de la métrique distance dans une carte du réseau	73
2.4	Illustration de la métrique bande passante dans une carte du réseau . . .	75
2.5	Illustration des coûts de transport dans une carte du réseau	78
2.6	Graphe de la fonction de coût avec $a=2$ and $b=1$	80
2.7	Graphe de la fonction de coût $a=2$ and $b=4$	81
2.8	Illustration des deux approches de calcul pour les coûts de transport et coûts OSPF	82
2.9	Illustration d'une matrice de proportions	86
2.10	Illustration d'une matrice de contraintes	88
2.11	Diagramme de flux de la procédure d'échange d'informations utilisant l'interface CINA entre opérateurs réseaux et gestionnaires de services applicatifs	88

3.1	Illustration de la plateforme CDN d'Akamai	97
3.2	Illustration de l'intégration de l'architecture CINA dans la plateforme CDN d'Akamai	98
3.3	Illustration de l'intégration de l'architecture CINA dans un tracker P2P	101
3.4	Illustration de l'intégration de l'architecture CINA dans un client P2P .	102
3.5	Illustration du schéma d'affectation des priorités aux fragments de données	106
3.6	Illustration des six premiers intervalles de temps du régime transitoire après la formation d'un groupe complet de huit noeuds.	109
3.7	Illustration des échanges de données en présence du pair de stabilité. Les fragments de données sont transmis en priorité à ce noeud particulier, identifié N_9 dans chaque groupe.	114
3.8	Illustration de deux constructions d'arborescence avec des tailles et des profondeurs d'arbre différentes.	117
4.1	Illustration de la modélisation d'un noeud d'un réseau à commutation de paquet par les files d'attentes	122
4.2	Les étapes d'exécutions des simulations analytiques	128
4.3	Graph du réseau RNP	130
4.4	Graph du réseau RNP	131
4.5	Le gain réalisé en termes d'utilisation moyenne du réseau des différentes techniques de redirection par rapport à une sélection aléatoire	133
4.6	La distribution des taux d'utilisation des liens du réseau des différentes techniques de redirection	133
4.7	Illustration de l'utilisation des liens du réseau durant une simulation pour un trafic initial à 40%	135
4.8	Illustration de l'utilisation des liens du réseau durant une simulation pour un trafic initial à 40%	135
4.9	Illustration de la moyenne et variance des taux d'utilisation des liens . .	135
4.10	Illustration des gains réalisés en utilisation et en latence des différentes techniques	136
4.11	Graph du coeur de réseau Orange Europe	139
4.12	Illustration du nombre de clients redirigés vers les différents serveurs inter-connectés avec le réseau	140
4.13	Le gain réalisé en termes d'utilisation moyenne du réseau des différentes techniques de redirection par rapport à une sélection aléatoire	142
4.14	La distribution des taux d'utilisation des liens du réseau des différentes techniques de redirection	142
4.15	Illustration de l'utilisation des liens du réseau durant une simulation pour un trafic initial à 40%	144
4.16	Illustration de l'utilisation des liens du réseau durant une simulation pour un trafic initial à 40%	144
4.17	La distribution des taux d'utilisation des liens du réseau des différentes techniques de sélection pour une taille d'arbre de 5	148

4.18	La distribution des taux d'utilisation des liens du réseau des différentes techniques de sélection pour une taille d'arbre de 10	148
4.19	La distribution des taux d'utilisation des liens du réseau des différentes techniques de sélection pour une taille d'arbre de 15	149
4.20	Illustration de la moyenne et de la variance des taux d'utilisation des liens pour une taille d'arbre de 5	149
4.21	Illustration de la moyenne et de la variance des taux d'utilisation des liens pour une taille d'arbre de 10	150
4.22	Illustration de la moyenne et de la variance des taux d'utilisation des liens pour une taille d'arbre de 15	150
4.23	Illustration des gains réalisés par rapport à la technique de sélection aléatoire pour différentes tailles d'arbres.	152
4.24	Illustration de la latence moyenne dans le réseau pour les différentes techniques	152
4.25	Illustration des échanges entre les entités fonctionnalités développées pour le prototype CDN	155
4.26	Illustration des étapes de fonctionnement du prototype à travers les interactions entre les différentes entités fonctionnelles	157
4.27	Illustration de la plateforme de tests avec les différentes entités du service applicatif	161
4.28	Débit du service applicatif du PID 1	163
4.29	Débit du service applicatif du PID 2	163
4.30	Débit du service applicatif du PID 3	163
4.31	Débit du service applicatif du PID 5	164
4.32	Débit du service applicatif du PID 7	164
4.33	Débit du service applicatif du PID 8	164
4.34	RTT moyen pour les sessions des utilisateurs du PID 1	166
4.35	RTT moyen pour les sessions des utilisateurs du PID 2	166
4.36	RTT moyen pour les sessions des utilisateurs du PID 3	166
4.37	RTT moyen pour les sessions des utilisateurs du PID 5	167
4.38	RTT moyen pour les sessions des utilisateurs du PID 7	167
4.39	RTT moyen pour les sessions des utilisateurs du PID 8	167
4.40	Répartition des taux d'utilisation des liens pour les différentes techniques de sélection	168

Liste des tableaux

1.1	Synthèse de différentes propositions académiques pour la problématique de gestion trafic par interface entre services applicatifs et réseaux . . .	40
2.1	Distance entre le client et les noeuds de distribution	73
2.2	Utilisation de la métrique bande passante pour une matrice de coûts . .	76
2.3	Comparaison entre les coûts OSPF et les coûts de transport	82
2.4	Comparaison des résultats de différentes fonctions de coût	83
2.5	Un exemple d'une matrice de coûts	84
2.6	Un exemple d'une matrice de redirection compacte	84
2.7	Exemple d'une matrice LU	90
2.8	Les paramètres du modèle	90
3.1	Les paramètres du modèle	105
3.2	La table de routage pour le régime périodique d'un groupe de taille 8 pour un cycle de taille 8	110
3.3	Comparaison des performances de notre modèle par rapport aux limites théoriques d'optimalités. Les groupes P2P ont une taille de 8 noeuds avec des capacités de transmission différentes et un nombre de groupes descendants différents	113
3.4	Comparaison des performances du modèle avec ou sans l'utilisation d'un pair de stabilité	115
4.1	Les paramètres du modèle analytique	124
4.2	Principaux paramètres des simulations	129
4.3	Principaux paramètres des simulations	137
4.4	Principaux paramètres des simulations	145
4.5	Paramètres des tests	159
4.6	Configuration des services CDN	159
4.7	Présentation des résultats des redirections des clients du CDN-A vers les serveurs localisés dans les PIDs 6 et 9	160
4.8	Présentation des résultats des redirections des clients du CDN-B vers les serveurs localisés dans les PIDs 9 et 10	160

Introduction

0.1 Contexte et problématique

Internet tel que nous le connaissons aujourd'hui est sujet à une polarisation autour d'une minorité de géants du web qui rivalisent d'inventivité pour développer de nouveaux services, créer de nouveaux contenus et constituer une base de clients sur laquelle ils comptent assoir leurs modèles économiques. Nous assistons littéralement à une explosion du trafic internet qui touchent principalement le trafic web avec une nette augmentation de la part de la vidéo dans ce trafic. Dans les conclusions des mesures de trafic inter-domaine qu'ils ont menées pendant deux années entre 2007 et 2009, Labovitz et al. [LIJM⁺10] ont cité une croissance annuelle du trafic de 45 %. Dans le même sens, une étude prévisionnelle de Cisco [Cis13] estime que le trafic internet dans les réseaux mobiles va être multiplié par un facteur de 13 à l'horizon de 2017.

Cette croissance soutenue du trafic est engendrée surtout par une forte appétence des clients pour les services proposés par les acteurs du net et assurée par l'explosion des terminaux connectés dont le contenu technologique a connu un bond remarquable. Les réseaux communautaires se posent désormais comme un relai garant de cette croissance, plus particulièrement après l'adoption des services vidéos dans leurs vitrines. Dans ce contexte, les géants de l'Internet ne se contentent plus d'exploiter les réseaux de communications mais déploient leurs propres systèmes et techniques de distribution de leurs données. On assiste ainsi à une croissance effrénée des réseaux de distribution de contenus, Content Delivery Networks (CDN), à l'augmentation de leurs capacités, aux déploiements de " Data Centers (DC), etc. Cette ligne directrice est motivée par un besoin fondamental en satisfaction client à travers une qualité d'expérience, Quality-of-Experience (QoE), acceptable, et qui ne peut être dissociée de la qualité de service, Quality-of-Service (QoS), pour l'acheminement des données.

En effet, toutes les études [LIJM09, LIJM⁺10, San11] concordent sur l'ascension fulgurante de certains acteurs comme Google. Alors qu'il ne figurait même pas dans le top 10 des acteurs leaders du trafic Internet en 2007, la croissance fracassante du trafic de Google le propulse en troisième position en 2009 avec environ 5.20% du trafic à son compte (le rachat de Youtube en 2006 a transformé la firme) [LIJM09]. Ces développements ont bouleversé un secteur où les principaux acteurs sont traditionnellement des

opérateurs de transit et de télécommunications. En 2010, Google prend des parts supplémentaires pour atteindre près de 7% du trafic internet mondial [LIJM⁺10]. D'autres acteurs comme Amazon, Facebook ou encore Netflix connaissent aussi une importante croissance de leur trafic qu'ils tentent de contrôler par l'utilisation de réseaux de distribution de contenus CDNs. Une étude de Sandvine [San11] sur le trafic Internet en Amérique du Nord a montré que Netflix et Youtube à eux seuls représente près de 35% de la totalité du trafic Internet aux heures de pleine charge.

Ce basculement des services aux client finaux, depuis les opérateurs de télécommunications vers les acteurs de l'Internet nécessite que les premiers (les opérateurs) gèrent ce basculement de valeur des services aux clients finaux (services aux particuliers ou B2C) vers les services aux parties tierces (services de gros ou B2B). Par ailleurs, il devient évident que le trafic généré par les CDNs doit bénéficier d'une attention particulière. Cette attention concerne à la fois les opérateurs réseaux qui sont responsables d'acheminer les données jusqu'aux utilisateurs, mais aussi des opérateurs et exploitants des CDNs qui sont soucieux d'offrir des services avec une meilleure qualité.

Le problème qui se pose dans ce contexte concerne la gestion de cette croissance de trafic. En effet, bien que ces acteurs soient inter-connectés entre eux, ils opèrent de façon indépendante et même opaque les uns vis-à-vis des autres. Au moment où l'amélioration de la QoS s'impose dans les réseaux d'opérateurs, où la QoE est facteur de différenciation entre les fournisseurs de services internet, la considérable croissance du trafic présentent de sérieux défis voire soucis à traiter par chacun. Les approches individualistes des différents acteurs concernés se résument en des techniques de gestion unilatérale du problème à travers les leviers à disposition accompagnée ou non par des tentatives d'inférence des politiques, ripostes et besoins des autres acteurs. Or, une gestion efficace des problèmes d'augmentation de trafic requiert plus de synergie entre les différents acteurs.

0.2 Motivations, approche, et défis

Il est tout à fait dans l'air du temps aujourd'hui de privilégier l'optimisation de l'existant aux investissements et engagements financiers. Cette approche s'applique bien à cette problématique de croissance du trafic. Les opérateurs réseaux ne peuvent continuer à investir colossalement dans le sur-dimensionnement de la capacité de leurs réseaux. Face à l'effondrement de leurs revenus et de la valeur bande passante du réseau (fixe et bientôt mobile), les opérateurs sont plus réticents à sur-dimensionner et à ainsi détruire la valeur de leur principal atout qui est le réseau. Certaines limites sont atteintes et une nouvelle approche pour traiter donc cette problématique de manière différente s'impose. Nous privilégions en conséquence une meilleure gestion des ressources existantes, qui est une voie plus favorable sur tous les plans. Elle permet de repousser les investissements, d'améliorer la rentabilité et d'optimiser le rendement.

Notre approche est d'apporter une réponse efficace et conjointe de part et d'autre en prenant en considération les besoins et les contraintes de chaque acteur. Dans le cadre de notre thèse, nous avons étudié l'intérêt de l'utilisation d'une interface de collaboration entre les acteurs responsables du transport et ceux responsables du service. Cette collaboration passe par le partage d'informations profitables à la bonne gestion du trafic généré dans le réseau et à l'amélioration de la QoE perçue par les utilisateurs. L'interface peut être utilisée aussi pour le partage de services entre les acteurs toujours dans la perspective d'améliorer les performances communes. Parmi les services qui peuvent être proposés par les opérateurs réseaux, nous pouvons citer le service de transmission multicast ou bien le service de cache dans le réseau.

0.3 Contributions

En partant du constat des contraintes actuelles en trafic et en besoin en qualité de service pour des applications orientées multimédia et interactivité, et en analysant les limites des techniques proposés dans l'état de l'art, nous avons amorcé nos travaux par la spécification d'un cadre d'action à travers une architecture fonctionnelle simple, adapté et évolutif. Dans la partie suivante de la thèse, nous nous sommes intéressés à l'amélioration d'une approche spécifique proposée par le groupe de travail Application Layer Traffic Optimisation (ALTO) à Internet Engineering Task Force (IETF) [ALT12]. Les travaux qui ont été effectués au sein de ce groupe de normalisation ont permis de définir un cadre fonctionnel et spécifier un protocole de communications [APY12] entre un service ALTO dans le réseau et un client applicatif ALTO. Nous avons poussé nos travaux dans l'étude de l'application de différentes techniques qui exploiteront le protocole ALTO pour permettre aux opérateurs réseaux d'exposer différents types d'informations aux fournisseurs de services. Notre architecture fonctionnelle, définie conjointement dans le cadre du projet européen ENVISION [UCoL09, MES⁺11a], a été développée pour offrir plus de fonctionnalités que ce que permette ALTO. Nous pouvons citer à titre d'exemple la réciprocité des échanges d'informations le serveur dans le réseau et le client hébergé par les applications mais aussi la possibilité de bénéficier aux applications de certains services réseaux comme la transmission multicast IP.

Notre contribution s'est focalisée par la suite sur la spécification de l'interface de collaboration qui est d'une conception simple issue des environnements web et internet plutôt que télécom. L'objectif est de cacher la complexité du réseau vis-à-vis des fournisseurs de services pour faciliter l'adoption de l'interface tout en préservant l'intégrité du savoir faire des opérateurs. Cette interface, qui adopte le protocole ALTO, dispose de caractéristiques qui reflètent le contexte de sa mise au point. Elle s'associe à la démarche "*best-effort*" en s'absolvant de contraintes légales d'obtention de résultats sans renier les objectifs d'optimisations qu'elle est sensée remplir. Par ailleurs, nous avons fait en sorte de disposer de canaux symétriques permettant aux opérateurs et aux acteurs Internet d'exposer des caractéristiques, besoins, contraintes et préférences.

La principale partie de nos travaux porte cependant sur les techniques d'optimisation de la gestion du trafic dans le réseau. Nous avons évoqué la polarisation de l'Internet autour de certains acteurs géants. Ceci démontre tout d'abord que les opérateurs acheminent une grande partie de trafic qu'ils ne maîtrisent pas mais aussi qu'une partie non négligeable de ce trafic proviendrait d'un groupe d'acteurs limités. Nous avons étudié en conséquence des techniques qui permettent à la fois d'optimiser l'acheminement des données dans le réseau mais aussi de réaliser ces opérations au niveau applicatif là où le contrôle est plus efficace et plus simple. Les premières études ont porté sur la collaboration passive des opérateurs réseaux qui mettent des informations à disposition des fournisseurs de services. Ces derniers sont responsables de leurs exploitations en vue de prendre des choix judicieux au niveau applicatif. La deuxième approche proposée avance une technique de collaboration active où les acteurs s'échangent des informations dans une logique coopérative. L'optimisation de la répartition du trafic dans le réseau est réalisée dans ce cas-ci par les opérateurs qui sont le dénominateur commun de tous les services et applications concernés. En effet, cette technique permet aux opérateurs de prendre des décisions par rapport à la globalité du trafic qui sera généré par les applications. Nous avons étudié par la suite le cadre d'utilisation de ces techniques pour deux cas d'applications en vogue, CDNs et Peer-to-Peer (P2P), et qui sont liés au trafic le plus problématique car exigeant toujours plus de bande passante.

Nous avons entrepris aussi la réalisation de simulations et expérimentations afin d'évaluer la performance de nos approches et nos propositions. Nous avons pu ainsi comparer différents types de métriques comme la distance ou la bande passante disponible employés dans nos algorithmes d'optimisation. Nous avons pu étudier la répartition du trafic dans le réseau par rapport aux choix d'approches, de métriques, etc.

Ces contributions sont valorisées à travers des articles de conférences, des contributions significatives au projet européen ENVISION, des dépôts de brevets et un chapitre de livre.

0.4 Organisation du rapport

Ce rapport est structuré en quatre chapitre. Le premier est consacré à l'étude des interfaces de communication entre réseaux et services. Cette partie d'état de l'art permet de comprendre l'évolution des approches de gestion du trafic à travers différentes contributions scientifiques et industrielles. Les modèles étudiés sont variés depuis des modèles opérateurs "*Walled Garden*" jusqu'aux modèles académiques fusionnant les dimensions réseaux et services. Ces études nous ont permis d'identifier les limites des approches proposées, l'inadéquation des cadres d'application ainsi que les problèmes de complexité et de rigidité qui ont freiné le développement de plusieurs propositions.

Dans le second chapitre, nous étudions l'utilité d'un moyen simple, efficace et adapté à travers un regard critique sur l'état de l'art. Nous avançons nos propositions à travers

une architecture fonctionnelle et adaptée aux contextes de la problématique. Nous présentons ensuite les spécifications de l'interface de communication et les atouts qu'elle apporte dans cette nouvelle vision que nous portons. Enfin, nous développons les techniques d'optimisation que nous proposons et détaillons les algorithmes employés.

Le troisième chapitre a été consacré à l'étude de l'application des approches proposées dans deux systèmes fréquemment utilisés aujourd'hui, à savoir les réseaux de distribution de contenus et les applications pair-à-pair. Nous analysons le cadre de mise en oeuvre des interfaces et techniques d'optimisation. Nous élaborons des mécanismes optimaux pour le transport des données dans les réseaux pair-à-pair spécialement spécifiés autour de notre approche de collaboration.

Nous portons finalement dans le dernier chapitre l'ensemble des travaux réalisés pour l'évaluation de nos propositions. Nous présentons le développement de notre modèle de simulation et sa validation. Nous analysons par la suite les résultats des différentes simulations que nous avons réalisées. Ces simulations sont complétées par des expérimentations avec des prototypes déployés dans un réseau physique. Les résultats sont présentés en dernière partie.

Enfin, le manuscrit se conclut par une synthèse de nos conclusions et les perspectives de travaux de recherche et d'exploitation auxquels ce rapport ouvre la voie.

Chapitre 1

Analyse des différentes interfaces de communications entre services et réseaux pour la gestion des flux de données

La séparation entre les couches de transport et les couches applicatives qui les utilisent pour communiquer a été induite par l'organisation en couches des modèles OSI Open Systems Interconnection (OSI) et IP. Cette organisation a fait l'objet de vastes travaux de recherche pour améliorer la gestion des flux de données au niveau du transport. En effet, d'un côté un besoin concret est apparu avec la diversification des services et applications. Il s'agit à titre d'exemple de garantir une certaine qualité de service à des flux spécifiques, de permettre l'utilisation de certaines ressources par d'autres flux, etc. D'un autre côté, l'augmentation continue de la charge de trafic dans le réseau impose le recours à des techniques de gestion plus sophistiquées. Dans un contexte où les réseaux sont de plus en plus marginalisés dans l'équation de transport, les marges de manoeuvres sont limitées à l'ingénierie du trafic, à l'utilisation de ressources de cache, etc. Bien que ces solutions pourraient apporter une certaine amélioration, non seulement elles reposent sur des contraintes (l'ingénierie de trafic est un processus très lourd et les serveurs de caches sont exploitables uniquement pour certains flux non chiffrés), mais aussi, elle n'offre pas la perspective d'une approche de gestion optimisée.

Dans ce cadre précis, il y a des travaux qui ont été entrepris en recherche, en normalisation et en projets mixtes afin de définir des interfaces de communication entre services et réseaux. Ces travaux avaient pour objectif d'apporter des éléments de réponses à des problèmes techniques ou bien à des besoins de performances. Nous procédons dans ce chapitre à l'étude et l'analyse de ces différents travaux pour cerner leurs spécificités et déterminer leurs utilités. Nous présenterons les interfaces et les techniques objet de leurs travaux, leurs intégration dans les architectures de services et de réseaux et exposerons les contraintes de leur mise en application.

Nous commencerons le chapitre par la présentation de l'architecture ALTO [ALT12] et des travaux dédiés. Cette approche est un pas important vers l'ouverture des réseaux sur les applications tierces exclusivement pour l'optimisation du transport des données. L'interface définie dans ce cadre a pour but de permettre l'exposition d'informations réseaux au bénéfice d'une meilleure gestion du transport des données dans le réseau à travers des choix mieux orientés au niveau applicatif. Dans la deuxième section, nous exposerons des travaux de recherche qui se sont intéressés à la résolution de cette problématique. Les interfaces sont parfois explicitement définies et reposent sur ALTO, parfois non. Des techniques d'optimisation sont proposées en ce sens pour exploiter ces échanges entre services et réseaux. Finalement, nous traiterons des efforts des organismes de normalisation sur le sujet. La finalité de leurs travaux est divergente de la nôtre. Ces systèmes définissent des processus pour contrôler le trafic dans le réseau pour des besoins de QoS des services de téléphonie, de TV par exemple. Les interfaces de communication permettent ainsi de définir le cadre de ce genre de contrôle. Nous assistons aussi à la spécification d'interfaces qui offrent la possibilité d'exploiter des fonctionnalités réseaux par des applications tierces mais rarement pour des besoins de performances ou d'optimisation. Les fonctionnalités les plus demandées concernent la messagerie, la localisation, la facturation et la qualité de service.

1.1 ALTO, l'interface de référence

Le groupe de travail ALTO [ALT12] de l'IETF travaille sur la standardisation d'une interface entre réseaux et applications qui utilise un protocole appelé ALTO [APY12]. Le groupe a pour objectif de concevoir et spécifier une interface qui permet aux applications d'avoir accès à diverses informations sur les réseaux de transport qu'elles utilisent. Ces informations devraient permettre aux applications de faire des choix plus appropriés quant à la gestion du trafic mais toujours au niveau applicatif. A titre d'exemple, choisir des pairs dans la même zone pour un client P2P réduira le trafic sur des liens d'interconnexion avec les autres zones et offrira en parallèle de meilleures performances. Le groupe de travail s'est intéressé plus particulièrement aux besoins des applications P2P au départ mais a intégré par la suite les problématiques de sélection de serveurs CDNs. En effet, le groupe de travail ALTO a permis de porter en standardisation à l'IETF des travaux qui ont été initiés par le projet A Portal for Proactive network Provider Participation for P2P (P4P). Nous présenterons dans les deux sections suivantes les travaux du projet P4P et ceux du groupe ALTO pour pouvoir aborder par la suite l'ensemble des travaux qui les ont repris comme modèles.

1.1.1 Le projet P4P

Le projet P4P [XKSY, XYK⁺08] pour "*A Portal for Proactive network Provider Participation for P2P*" propose un système pour répondre à la forte croissance du trafic des applications pair-à-pair ou P2P. En particulier, le projet P4P fait valoir qu'il existe un problème dans la gestion et le contrôle des réseaux de transport en ignorant les applications. Il considère que les nouvelles applications peuvent avoir une grande

flexibilité dans la gestion de leurs communications, et par conséquent, elles devraient être intégrées dans la chaîne de gestion et de contrôle du trafic.

Dans ce contexte, le projet P4P considère que si les applications et utilisateurs finaux doivent participer à l'optimisation des ressources réseau, les opérateurs de ces réseaux ne peuvent pas continuer à garder une opacité absolue. Au contraire, les opérateurs réseaux devraient être amenés à exposer un certain nombre d'informations telles que leurs politiques, leurs topologies, etc. En conséquence, le projet propose un moyen de communications entre réseaux et applications.

1.1.1.1 Le système P4P

La proposition du projet P4P se base sur un système simple et flexible qui permet aux opérateurs réseaux de fournir des informations ou recommandations explicites aux applications P2P. Le système intervient sur deux plans différents : le plan de commande et le plan de données.

Dans le plan de commande, P4P introduit le module iTrackers. Ce module fournit la fonction portail qui permet aux applications P2P de communiquer avec les opérateurs réseau. En effet, chaque opérateur réseau maintient un ou plusieurs iTracker dans son réseau. Un client P2P obtient l'adresse IP de l'iTracker de son fournisseur local par requête DNS. Les techniques standard permettent d'assurer la redondance des iTrackers dans un domaine particulier pour améliorer la tolérance aux pannes ou assurer une meilleure flexibilité d'évolution.

L'iTracker qui implémente le portail pour les applications P2P, offre trois types d'informations relatives au réseau dont il est représentant. Il s'agit de l'état du réseau et sa topologie, les politiques et directives de l'opérateur et ses capacités.

Dans le plan de données, P4P permet aux routeurs d'intervenir afin de fournir un retour (une boucle rétroactive) aux applications P2P pour permettre une utilisation plus efficace des ressources réseau. Plus précisément, les routeurs peuvent marquer les bits Explicit Congestion Notification (ECN) [RFB01] (ou un champ dans une en-tête P2P) pour explicitement désigner les sessions. Les clients finaux se chargent ensuite d'ajuster leurs débits en conséquence.

1.1.1.2 Le plan de commande du système P4P

Le système défini par le plan de commande de P4P fournit une interface de communication entre les entités applicatives clients ou trackers P2P et les réseaux sous-jacents. La figure 1.1 illustre les entités potentielles du système : l'iTracker appartenant au réseau, le appTracker de l'application P2P, et les clients P2P. Dans une configuration particulière (par exemple une application P2P totalement distribuée) certaines entités pourraient être ignorées (les appTrackers par exemple). P4P ne précise pas les flux d'échanges d'informations, mais fournit plutôt un cadre de communication avec des messages de contrôle encodés en XML pour assurer une meilleure extensibilité.

L'élément clé du système P4P est le iTracker. Ce module fournit trois interfaces qui permettent aux applications P2P d'interroger le réseau de transport.

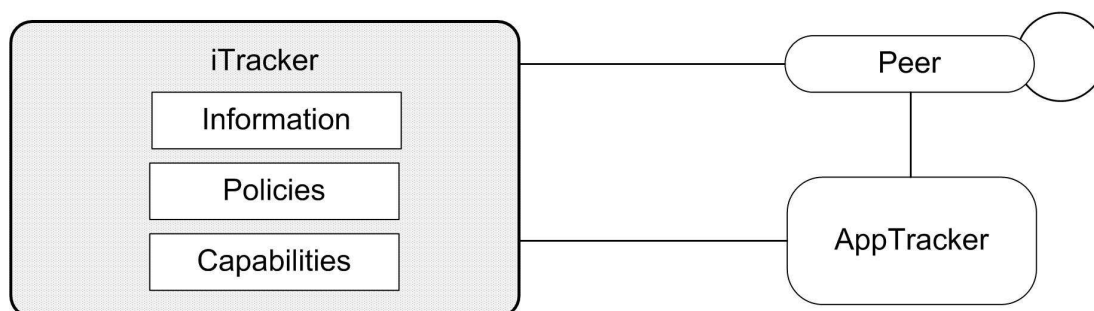


FIGURE 1.1 – Illustration de l'architecture du système de contrôle P4P

L'interface "*Information*" permet d'exposer des informations sur le réseau relatives à sa topologie et son état. Plus précisément, étant donné une requête sur une adresse IP appartenant au réseau, l'interface permet de localiser l'adresse en question en renvoyant un triplet (ASID, PID, LOC) :

- L'ASID est l'identifiant du réseau (par exemple, son numéro Autonomous System (AS))
- Le PID est un identifiant opaque affecté à un groupe de noeuds du réseau. Il est utilisé pour préserver la confidentialité des informations relatives à la topologie des opérateurs réseaux en présentant des vues moins fines du réseau.
- Le LOC fournit des coordonnées virtuelles ou géographiques de l'adresse IP. Il peut être utilisé pour déterminer la proximité des entités dans le réseau, ce qui peut être utile dans le choix des pairs avec qui on s'apprête à échanger des données.

Lors de l'envoi d'une requête d'information, un pair peut éventuellement transmettre l'identifiant du réseau de pairs auquel il est connecté (par exemple, les informations de hachage d'un torrent). Le iTracker peut garder la trace des pairs et des réseaux auxquels ils sont connectés.

L'interface "*Policies*" permet, d'obtenir des informations sur la politique de l'opérateur réseau. Ces informations permettent de spécifier comment un fournisseur de réseau souhaite que ses ressources soient utilisées. A titre d'exemple, la politique de l'opérateur peut énoncer le ratio acceptable entre les volumes de trafic entrant et sortant d'une zone ou bien les directives d'utilisation suivant les tranches horaires pour éviter d'utiliser des liens qui sont congestionnés pendant les heures de forte charge sur le réseau.

L'interface "*Capabilities*" permet aux applications P2P, pairs ou appTrackers, de demander des informations sur les capacités et ressources du réseau. Par exemple, un opérateur réseau peut fournir certains services comme des serveurs de cache dans le réseau. Un appTracker peut demander de l'iTrackers l'utilisation de ces serveurs en tant que pairs afin d'accélérer la distribution des données de l'application P2P.

1.1.2 Le protocole ALTO

ALTO définit une interface de communication entre un client et un serveur. Le protocole adopte un design RESTful [Fie00] qui est un modèle de service web simple et efficace reposant sur le concept de requêtes/réponses sans sauvegarde d'état. Le protocole est basé sur le protocole standard HTTP afin d'assurer une conformité avec les infrastructures HTTP largement déployées et bien maîtrisées. Les requêtes et réponses HTTP sont formatées suivant le schéma JSON. La collaboration nécessite l'implémentation d'un client ALTO par le service applicatif et d'un serveur ALTO par le réseau (voir figure 1.2).

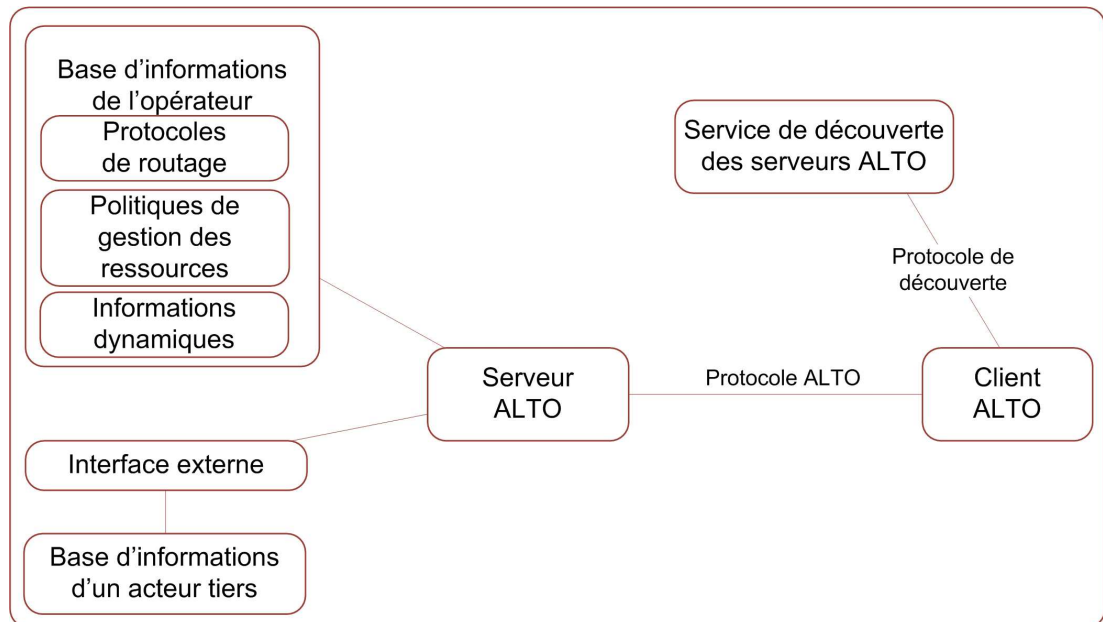


FIGURE 1.2 – Illustration de l'architecture fonctionnelle d'ALTO

Ce modèle client-serveur permet au service applicatif d'envoyer des requêtes au réseau pour demander des informations qui guiderait ses choix de sélection d'un noeud de distribution ou bien d'un groupes de pairs pour chaque client. Les informations fournies par les réseaux peuvent provenir des protocoles de routage, des politiques de provision de ressources de l'opérateur, des informations de surveillance du réseau, etc. Elles sont ensuite traitées et structurées sous la forme de services ALTO (voir figure 1.3). Nous nous intéressons particulièrement aux deux services principaux qui fournissent la carte du réseau, "*Network Map*" et la matrice des coûts, "*Cost Map*".

1.1.2.1 La carte du réseau

La carte du réseau fournit un ensemble d'emplacements réseau définissant des groupes appelés PIDs pour Provider-defined network location IDentifiers (PID). La politique de

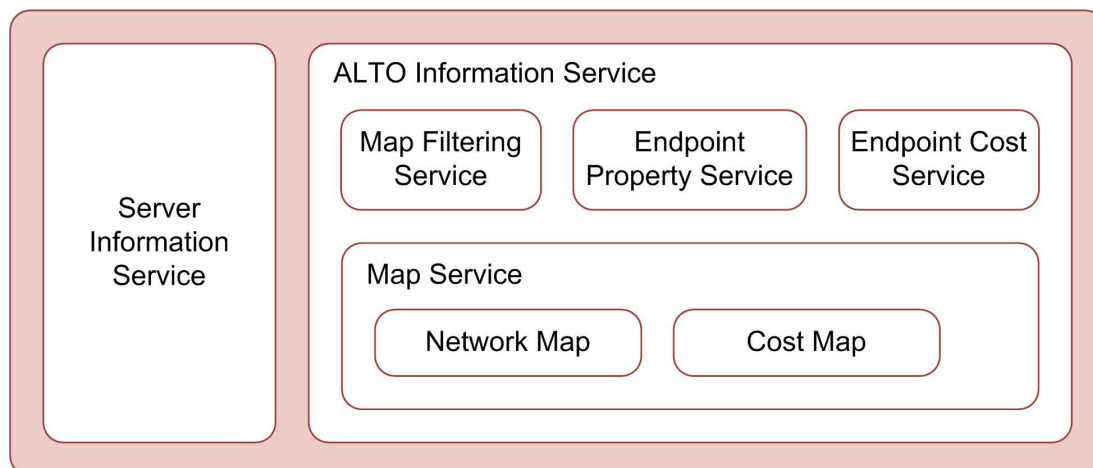


FIGURE 1.3 – Illustration des services ALTO disponibles à travers le protocole ALTO

spécification de ces groupes peut être basée sur différents critères comme la proximité géographique, la localisation dans le réseau, le type de réseau d'accès, le type de terminal, etc. Chaque PID est défini par l'ensemble des adresses IP, des adresses de réseaux, des identifiants de systèmes autonomes, etc, qu'il incorpore. La figure 1.4 illustre une carte réseau spécifiant trois PIDs avec les adresses réseaux qui les définissent. Les PIDs 1 et 2 représentent des sous-réseaux alors que le PID 3 représente le reste de l'Internet (toute adresse IP qui n'appartient pas aux PIDs 1 et 2). La localisation des entités ap-

```

{
  "map-vtag" : "225006996",
  "map" : {
    "PID1":{"ipv4" : ["172.10.10.0/24","172.11.100.0/24"]},
    "PID2":{"ipv4" : ["172.20.10.0/24","172.20.12.0/24"]},
    "PID3":{"ipv4" : ["0.0.0.0/0"], "ipv6" : [ "::/0" ]}
  }
}
  
```

FIGURE 1.4 – La figure illustre un exemple d'une matrice de coûts entre 3 PIDs.

plicatives dans cette carte du réseau fournie par l'opérateur est effectuée par recherche d'appartenance de leurs adresses IP à l'un des PID (voir figure 1.5). Les opérateurs ré-

seaux choisissent le niveau de granularité des informations qu'ils exposent dans la carte du réseau, par exemple une granularité du niveau des DSLAMs, des routeurs régionaux, des routeurs de transit, etc, en accord avec leurs politiques et/ou accords avec les fournisseurs de services. L'intérêt indéniable de cette carte du réseau est de fournir aux services applicatifs le moyen de localiser leurs entités dans le réseau suivant une convention commune avec les opérateurs réseaux sans que ces derniers n'aient à divulguer des informations confidentielles ou critiques comme leurs topologies et infrastructures.

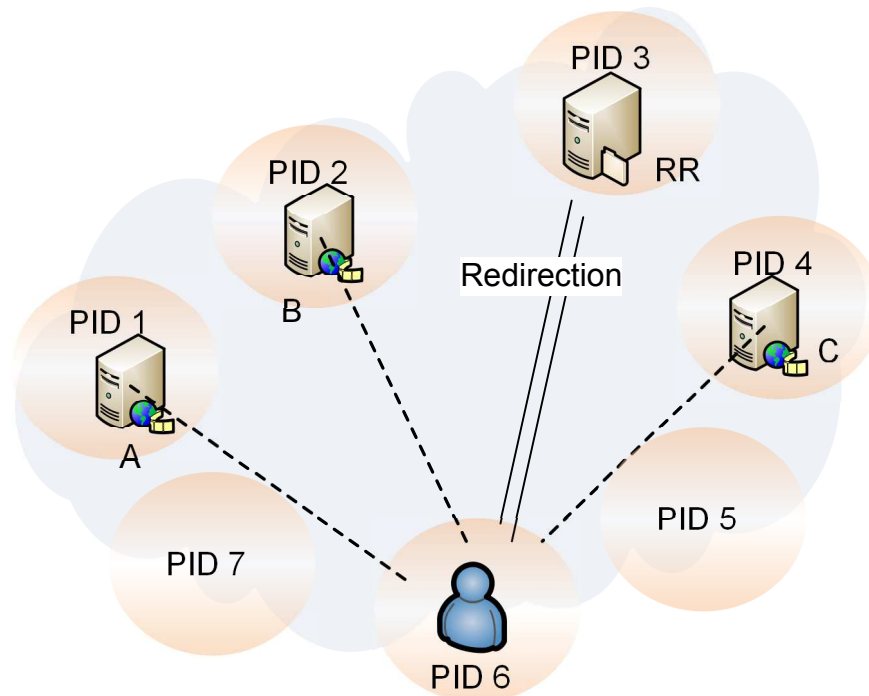


FIGURE 1.5 – La figure illustre un exemple d'une matrice de coûts entre 3 PIDs.

1.1.2.2 La matrice des coûts

La matrice des coûts est un autre service proposé par les opérateurs réseaux pour permettre aux fournisseurs de service de déterminer le coût d'acheminement du trafic réseau entre deux entités applicatives du service. Cette matrice reprend les PIDs définis par la carte du réseau. Les coûts sont spécifiés de bout en bout entre les différents PIDs et non entre les entités applicatives. A charge du fournisseur de service de déterminer la localisation de ces entités en exploitant la carte du réseau. Les coûts spécifiés par la matrice peuvent exposer des valeurs concrètes de certaines métriques réelles comme la distance au sens réseau entre deux PIDs (c-à-d le nombre d'équipements ou liens réseaux), la latence,... ou bien des valeurs abstraites qui sont calculées par l'opérateur à partir d'une ou plusieurs métriques, de sa politique, de ses accords avec les fournisseurs de service et du contexte de cette collaboration (nature du service, nature de la

collaboration, etc). Les opérateurs réseaux choisissent la nature des informations qu'ils désirent exposer aux services applicatifs et la façon de le faire. Ils utilisent pour cela l'attribut *type* pour spécifier le type d'informations et l'attribut *mode* pour spécifier la façon de les interpréter (par exemple numérique, ordinal, etc). La figure 1.6 illustre une matrice encodée en JSON qui expose des coûts de transport abstraits entre 3 PIDs.

```
{
  "cost-mode" : "numerical",
  "cost-type" : "routingcost",
  "map-vtag" : "225006996",
  "map" : {
    "PID1": {"PID1":1, "PID2":5, "PID3":5},
    "PID2": {"PID1":5, "PID2":1, "PID3":10},
    "PID3": {"PID1":5, "PID2":15, "PID3":1}
  }
}
```

FIGURE 1.6 – La matrice de coûts expose des coûts de transport entre 3 PIDs. L'attribut *type* spécifie que les valeurs doivent être interprétées comme des coûts abstraits d'acheminement du trafic entre les PIDs. L'attribut *mode* spécifie que les valeurs sont numériques.

1.2 Les interfaces proposées dans la littérature

Divers travaux de recherche académique ont décelé l'importance d'aligner les contextes du transport dans le réseau et du trafic généré par les applications. Ils ont contribué à spécifier des interfaces et des techniques capables d'améliorer les rendements à différents niveaux. Nous aborderons dans les sections suivantes différentes approches proposées à travers deux visions différentes. La première concentre les travaux sur l'optimisation du transport à travers l'exploitation d'informations sur les réseaux. La deuxième préconise l'alignement de l'ingénierie de trafic sur les contraintes applicatives.

1.2.1 Des informations réseau pour une meilleure gestion du trafic au niveau applicatif

Qu'il s'agisse d'applications, P2P, CDN ou d'autres natures, rapprocher les sources et les destinations des données et choisir un chemin réseau plus adapté est une stratégie qui a été largement étudiée dans la littérature. Les avantages escomptés sont justifiés par une meilleure qualité de service offerte par les réseaux de transport avec un impact positif sur les services. Nous présenterons dans les sections suivantes certains travaux académiques qui ont été entrepris dans la voie d'un meilleur routage du trafic au niveau applicatif grâce à des informations relatives aux réseaux de transport.

1.2.1.1 PaDIS

Le système PaDIS, "*Provider aided Distance Information System*", proposé dans [PFA⁺10], a pour objectif d'utiliser des informations fournies par les opérateurs réseaux aux applications *overlay* dans le but de réaliser une meilleure répartition du trafic dans le réseau. Plus précisément, c'est le cas des fournisseurs de services CDN qui a été étudié. Le système supporte des informations de type distance, latence ou bande passante.

Les auteurs commencent par démontrer, en analysant des traces de trafic, que la plupart des requêtes d'accès aux données (environ 70%) des utilisateurs des principaux opérateurs européens sont disponibles dans divers sites. Cette diversité justifie le besoin de profiter des informations pertinentes dont disposent les opérateurs réseaux afin de réaliser un choix approprié concernant la source à mobiliser pour servir les requêtes.

L'architecture du système PaDIS (illustrée dans la figure 1.7) est très similaire à celle d'ALTO. En effet, PaDIS est présenté comme un service proposé et géré par les opérateurs de réseaux. Les différentes applications, P2P, CDN, serveurs DNS, etc, utilisent le service PaDIS pour demander à l'opérateur de déterminer la meilleure source de données à utiliser pour servir un client spécifique. Il s'agit d'un ordonnancement d'adresses IP de la part de l'opérateur suivant les informations dont il dispose (distance, latence, etc).

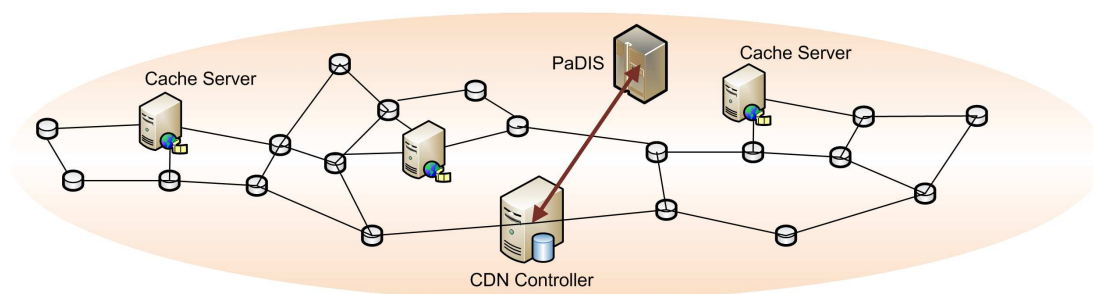


FIGURE 1.7 – Illustration de l'intégration du système PaDIS

1.2.1.2 CaTE

CaTE, pour "*Content-aware Traffic Engineering*" [FPS⁺12], est une proposition pour exploiter les informations disponibles sur le transport dans la gestion du trafic des CDNs. CaTE s'appuie sur l'assomption qu'en sélectionnant le noeud de distribution le plus approprié parmi ceux qui sont capables de satisfaire la requête d'un client, la qualité du transport des données dans le réseau peut être améliorée car le chemin offre une meilleure disponibilité. Pour permettre aux CDNs de bénéficier d'informations sur les conditions de transport dans leur sélection de serveurs, le système CaTE a été conçu comme un module capable de décider des chemins réseaux à emprunter par les flux de données.

Ce module est implémenté par les opérateurs réseaux pour s'interfacer avec leurs vis-à-vis de services. En ce sens, CaTE est très similaire à PaDIS. L'interface ALTO est

évoquée comme possible candidate aux besoins d'échanges d'informations entre réseaux et services.

Le positionnement de CaTE par rapport au Traffic Engineering En parallèle, les auteurs de CaTE s'opposent à toute intervention dans les procédures de "*Traffic Engineering*" (TE) des réseaux. Ils considèrent que le TE est basé sur des règles de répartition de poids bien étudiées pour aboutir sur une politique de routage saine et adaptée aux variations de la matrice de trafic dans le réseau. Pour éviter les micro-boucles lors de la convergence du protocole IGP [FB07], il est de pratique courante de seulement modifier un nombre réduit de poids [FT02]. Pour limiter ce nombre tout en proposant une politique de routage corrélée aux contraintes de trafic, les ingénieurs de TE s'appuient sur des matrices de trafic calculées sur de longues périodes et testées hors des systèmes opérationnels. En conséquence, CaTE est positionné comme une solution plus adaptée pour faire du TE au niveau applicatif sans se trouver soumis aux contraintes des réseaux opérationnels.

Le modèle de CaTE Dans CaTE, le réseau est modélisé par un graphe orienté $G(V, E)$ où V est l'ensemble de noeuds et E est l'ensemble des liens. Un flux dont l'origine est O et la destination est D , noté f_{od} , est composé de l'ensemble des données qui entrent dans le réseau au noeud $o \in V$ et sortent par le noeud $d \in V$. Le trafic sur un lien est la somme de tous les flux OD qui traversent le lien. La relation entre les flux de trafic et les liens du réseau est exprimée par la matrice de routage A . Cette matrice, de taille $|E| \times |V|^2$, est une matrice booléenne dont chaque élément indique si le flux OD traverse le lien ou non. y a été défini comme un vecteur de taille $|E|$ pour représenter le trafic sur les liens du réseau. x a été défini comme un vecteur de taille $|V|^2$ pour représenter l'ensemble des flux de trafic OD . De ces notations, les auteurs ont défini l'équation suivante $y = Ax$.

L'objectif du modèle L'approche du système CaTE est d'utiliser les informations réseaux afin d'ajuster la demande de trafic représentée par le vecteur x . En conséquence, étant donnée une matrice de routage A , le vecteur de charge sur les liens du réseau y peut être amélioré. Pour améliorer la demande de trafic, les auteurs définissent $x = x_r + x_s$ avec x_r le vecteur du trafic qui peut être ajusté et x_s le vecteur du trafic qui ne peut être changé.

L'objectif est de déterminer le vecteur optimal x' , étant donné une topologie G , une matrice de routage A et un vecteur trafic initial x , pour minimiser le taux d'utilisation maximal parmi tous les liens du réseau.

1.2.1.3 DONAR

Le système DONAR [WJFR10] est un système distribué qui a été étudié pour concevoir une technique de redirection des requêtes clients de façon efficace, sûre et performante. Le système repose sur le caractère distribué de la solution. C'est un système de

sélection décentralisé qui tend à optimiser le processus de sélection de noeuds de distribution en prenant en considération des informations sur le réseau (la localisation du client, le coût du transport, etc), la charge des serveurs, et les politiques des fournisseurs de services.

Le modèle du système DONAR Ce système distribué est constitué par une couche composée par plusieurs noeuds appelés "*Mapping Nodes*" qui sont responsables de la gestion des requêtes et de leur redirection vers les différents noeuds de distribution appelés "*Replicas Servers*". Ce système s'interpose entre les clients et les noeuds de distribution des fournisseurs de services Cloud ou CDN comme illustré par la figure 1.8 [WJFR10].

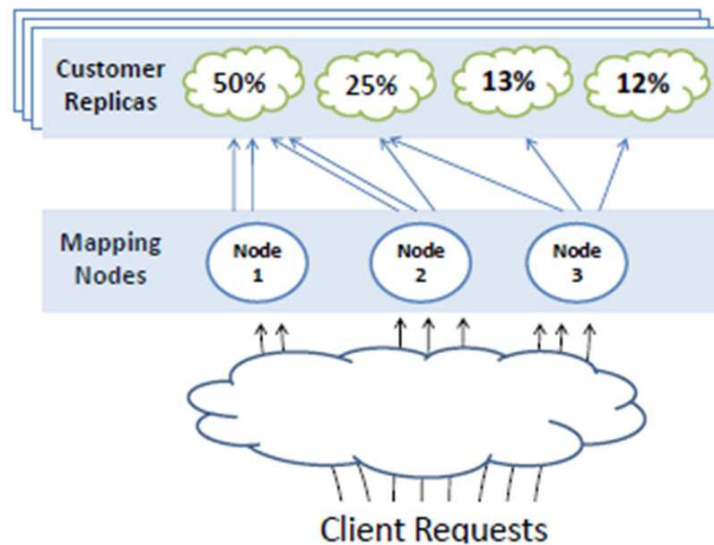


FIGURE 1.8 – Illustration du modèle système DONAR

Pour réaliser des redirections pouvant atteindre des résultats performants, chaque noeud de redirection doit disposer d'informations sur le service et le réseau pour s'adapter aux conditions changeantes. DONAR critique l'approche la plus simple qui consiste à avoir un organe central de coordination pour la collecte et le traitement des informations. En effet, les auteurs considèrent qu'un organe central présente un point critique car en cas de défaillance ou d'attaque ciblée, le système est inopérable.

La dimension réseau du système DONAR La politique de sélection de noeud de distribution du système DONAR repose sur plusieurs critères dont un critère réseau. Dans cette optique, le chemin réseau emprunté est considéré important car il a un impact significatif sur la performance de bout en bout. Les auteurs précisent par exemple que les contenus Web interactifs nécessitent une latence, exprimée en "*Round Trip Time*"

(RTT), limitée dans le réseau. En contraste, les services de transfert de gros volumes de données par exemple nécessitent quant à eux un bon débit de transfert dans le réseau.

L'algorithme de DONAR utilise en conséquence une fonction de coût abstraite $cost(c, i)$ pour s'adapter à ces différentes contraintes tout en caractérisant la performance du chemin réseau entre un client particulier et le noeud de distribution ciblé. Cette fonction est considérée comme très flexible. Elle permettrait par exemple d'optimiser les temps de latence du service en utilisant un coût dépendant uniquement du RTT du réseau. Les valeurs du RTT peuvent être mesurées directement dans le réseau ou bien estimées à partir d'un système de coordonnées implanté dans le réseau.

Le modèle d'optimisation de DONAR Pour formuler le problème de sélection de noeuds de distribution, les auteurs ont introduit un modèle réseau dans lequel C est l'ensemble des clients et I est l'ensemble des serveurs. R_{ci} a été définie comme la proportion de trafic du client c acheminée au serveur i . Ces valeurs font l'objet du problème d'optimisation dont la performance globale est calculée dans l'équation 1.1.

$$Perf = \sum_{c \in C} \sum_{i \in I} R_{ci} \times cost(c, i) \quad (1.1)$$

L'algorithme est soumis à des contraintes de services. Les auteurs proposent une résolution distribuée de l'algorithme offrant une plus grande indépendance aux noeuds de redirection introduits par le système, même s'ils doivent conserver des canaux de communications entre eux pour échanger des données vitales pour la résolution des calculs.

1.2.1.4 Le projet NAPA-WINE

Le projet "*Network-Aware P2P-TV Application over Wise Networks*"¹ (NAPA-WINE) est un projet européen qui a pour objectif de développer des techniques performantes pour la diffusion de la TV dans des réseaux P2P. Le projet propose une architecture où les clients de leur système P2P-TV exploitent des mesures dans le réseau (RTT, estimation distance, etc) pour réduire leur empreinte sur le réseau. Le projet propose aussi d'exploiter des services de type ALTO quand ils sont offerts par les opérateurs réseaux. L'objectif étant d'optimiser la performance du système P2P-TV tout en réduisant la charge globale de trafic exercée sur le réseau. En effet, le système P2P-TV compte exploiter les informations collectées à travers le protocole ALTO pour optimiser le processus de sélection de pairs. Par exemple, les clients P2P seront connectés en grande partie à des clients de leur voisinage afin de réduire les distances entre eux et améliorer les temps de réponse.

L'architecture de NAPA-WINE Le système P2P-TV de NAPA-WINE [BLM⁺11, ABB⁺10] repose à la fois sur ses estimations propres effectuées par ses clients et sur les

1. NAPA-WINE (<http://www.napa-wine.eu>) est un projet européen d'une durée de trois ans soutenu par la Commission Européenne dans le cadre du programme 7 (FP7-ICT-2007-1, Networked Media, grant No. 214412)

informations qu'il est susceptible de recevoir des opérateurs réseaux ou acteurs tiers pour optimiser son fonctionnement. Ces fonctions sont assurées par les modules *"Monitoring"* et *"Repository"* de l'architecture globale du projet illustrée dans la figure 1.9 [BLM⁺11]. Le module *"User"* est responsable des gestions de interfaces avec l'utilisateur. Le module *"Overlay"* s'occupe des fonctionnalités nécessaires au réseau P2P comme la sélection de pairs, etc. Finalement le *"Scheduler"* est lui responsable des transmissions et réceptions des *chunks*.

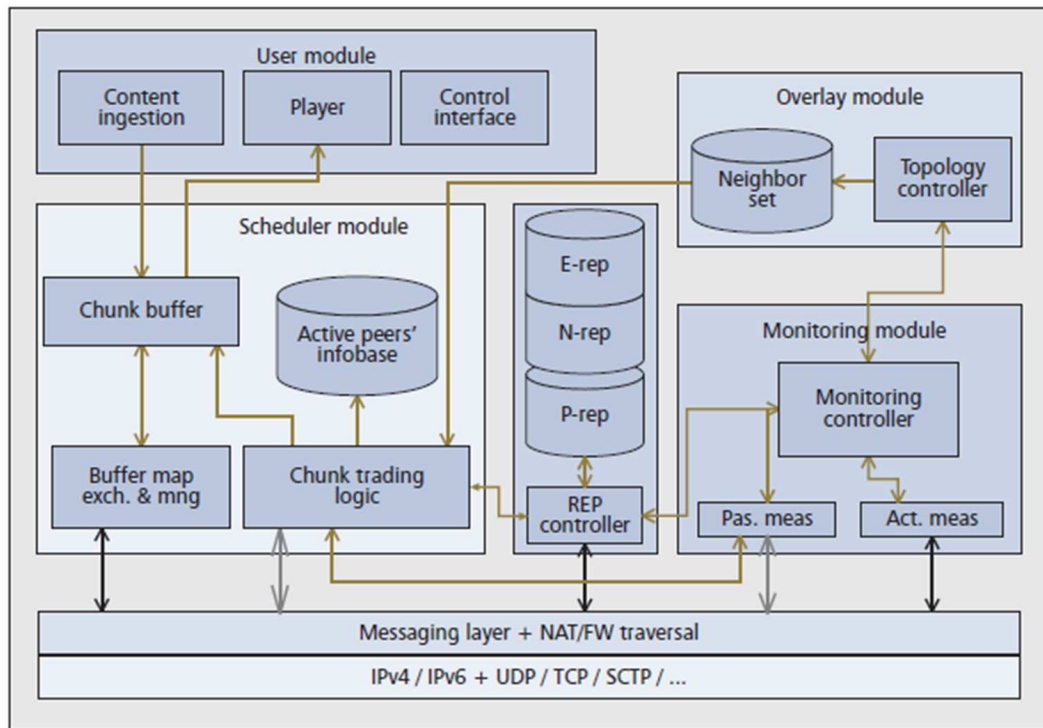


FIGURE 1.9 – Illustration de l'architecture globale du projet NAPA-WINE

Les modules Monitoring et Repository Le module "Monitoring" a été conçu pour recueillir et partager des informations utiles pour la construction et la gestion de la topologie du réseau P2P et l'amélioration des processus de d'échanges de *chunks* vidéo. L'objectif principal est d'assurer la disponibilité d'informations à jour sur la qualité des chemins de bout en bout entre les pairs et d'en déduire des informations sur les ressources disponibles du réseau. Plusieurs métriques réseaux sont estimées comme le RTT, le taux de perte de paquets, la bande passante disponible, le nombre d'équipements sur le chemin, etc.

Les modules "Repository" de l'architecture NAPA-WINE est composé de bases de données globales contenant des informations utiles pour les processus de décision relatif à la sélection des pairs pour chaque client ou à la transmission des *chunks*. Ces bases

de données sont un élément clé pour permettre au système P2P-TV de NAPA-WINE d'être capable de prendre en considération le contexte du réseau dans sa gestion. On distingue les trois suivantes :

- P-REP (*peer repository*) : c'est une base de données distribuée qui sauvegarde des informations sur les peers et les performances de bout en bout des chemins réseau qu'ils utilisent.
- N-REP (*network repository*) : cette data-base sauvegarde des informations globales sur le réseau à partir des collectes effectuées dans des P-REP.
- E-REP (*external repository*) : cette data-base est dédiée à sauvegarder des informations de source externe sur le réseau. La conception de ce module est destiné aux opérateurs réseaux qui acceptent de lui fournir des informations relatives à leurs réseaux comme par exemple les graphes des *Autonomous System*, les points de *peering*, les performances des chemins réseau, les tables de routage, la topologie, etc. le E-REP peut également stocker des informations relatives uniquement à un seul domaine. Pour assurer la protection des informations exposées par les opérateurs réseaux, le projet NAPA-WINE propose que le module E-REP soit installé et géré directement par les opérateurs réseau. Par ailleurs le E-REP dispose également d'une interface vers les serveurs ALTO.

1.2.1.5 I-CDN

I-CDN pour "ISP operated Content Delivery Network" [KMK⁺10] est une proposition de déploiement de systèmes de distribution de contenus CDNs dans les réseaux d'opérateurs. L'objectif étant de réduire la charge de trafic sur les réseaux tout en améliorant la QoE des services. Dans l'article, cet objectif serait réalisable en exploitant les informations valorisantes dont disposent les opérateurs de réseaux.

Ces informations serviront à optimiser les divers phases relatives au déploiement et fonctionnement des CDNs. Ainsi le positionnement des divers noeuds de distribution a été étudié en fonction de la topologie du réseau et la répartition des clients. Le processus de sélection d'un noeud de distribution pour servir un client est lui aussi assujetti aux informations à disposition des opérateurs.

1.2.1.6 Les systèmes de coordonnées virtuelles

Différents travaux de recherche ont porté sur la spécification de techniques capables d'aboutir à des concentrations locales de trafic. Ces techniques reposent sur la minimisation de la distance entre l'utilisateur et les divers sources de données capables de satisfaire ses requêtes (les pairs pour les systèmes P2P ou les noeuds de distribution pour les CDNs à titre d'exemple). L'objectif est d'emprunter un chemin réseau plus court afin d'améliorer la performance de transport (comme la latence par exemple) et éviter des chemins de longue distance empruntant des liens de transit éventuellement.

Une classe d'algorithmes [Bak05, BP08, HAOF05, RHKS02, BJZP09, MDJ11] proposée dans la littérature repose sur le concept des approximations topologiques. Le

concept se traduit par un système de coordonnées virtuelles capable de déterminer la localisation des entités applicatives et en conséquence d'estimer les distances entre elles. Dans [BP08], les auteurs proposent une approche de sélection de noeud de distribution CDN basée sur des informations relatives à la charge et à la proximité. Les informations de charge comprennent des données sur le taux d'utilisation des serveurs. Les informations de proximité se basent sur des coordonnées virtuelles dans le réseau, calculées au niveau applicatifs pour estimer la distance entre le client et les différents serveurs. Dans l'architecture qu'ils proposent, les auteurs utilisent un proxy qui est questionné par les clients afin de déterminer le noeud de distribution le plus approprié. Bakiras [Bak05] propose un réseau artificiel composé d'objets, de serveurs et de clients. Il définit une fonction de coût pour calculer le coût d'utilisation d'un serveur particulier par un client à partir de trois paramètres : la quantité de trafic, la capacité du serveur et les temps de traitement et de propagation. Dans sa proposition, le délai de propagation est calculé à partir d'estimations de latence (messages *ping*) entre des agents de redirection. Les différentes entités applicatives sont ensuite localisées par rapport à ces agents. Les travaux réalisés dans [RHKS02] proposent une technique de regroupement distribuée où les clients sont situés dans des zones. Un ensemble de noeuds de référence répartis sur Internet est utilisé pour localiser les clients en estimant leurs distances, calculées à partir des RTTs, par rapport aux noeuds de référence. Ces distances permettent de situer chaque client dans une zone. L'approche se base ensuite sur le système de coordonnées virtuelles pour déterminer le serveur le plus proche de chaque client. D'une manière similaire, les auteurs de cet article [HAOF05] utilisent des repères dans le réseau pour définir une topologie virtuelle permettant de situer les serveurs et regrouper ceux à proximité d'un repère. Les clients sont aussi localisés par rapport à ces repères et donc par rapport aux serveurs. Dans un autre article [BJZP09], les auteurs ont développé une technique de répartition de charge entre les noeuds de distribution d'un CDN qui va au delà de la simple variable de charge. Ils proposent d'intégrer des informations relatives à la topologie du réseau dans le système afin d'améliorer l'utilisation du réseau et la QoE des services CDNs. La topologie du réseau est définie à travers l'utilisation de points de référence dans le réseau qui servent de moyens pour la création de zones. Des mesures de RTT entre ces points de référence permettent par la suite d'obtenir des estimations de distance ou de performance de transport entre ces zones. Finalement, dans [MDJ11], les auteurs utilisent un outil de mesure appelé ABET pour "*Available Bandwidth Estimation Tool*" pour estimer la bande passante disponible de bout-en-bout entre les serveurs et les clients. Ces mesures sont effectuées entre un équipement de réseau implémentant le système ABET, un routeur par exemple, et l'ensemble des serveurs. Ces mesures sont ensuite transmises au serveur de gestion pour aider dans le choix du noeud de distribution le plus approprié.

Une approche différente au problème consiste à chercher à minimiser des coûts spécifiques et en particulier des coûts de transport dans le réseau ou des coûts d'utilisation des serveurs [CE10, WLCM11, WJFR10, FPS⁺12]. Carlsson et Eager [CE10] ont défini dans leurs travaux une fonction de coût à partir de la somme des coûts d'utilisation des serveurs et des coûts d'utilisation du réseau. Ils ont proposé un modèle avec une

topologie de N serveurs et M zones réseau, avec des serveurs locaux annexés aux zones. Pour caractériser le coût du réseau, ils l'ont défini en tant que variable binaire ; égale à 0 lorsque le client est redirigé vers le serveur local de sa zone et égale à 1 autrement. Dans d'autres travaux [WLCM11], une stratégie optimale pour la redirection des requêtes clients est proposée. Les auteurs ont défini un modèle caractérisant les clients, les serveurs et les services. Ils intègrent le délai de latence entre serveurs et clients comme variable dans le calcul du profit du CDN : un délai plus grand résulte en un profit inférieur. D'autres approches que nous avons précédemment évoquées [FPS⁺12, WJFR10] sont aussi basées sur la minimisation de certains coûts relatifs aux réseaux.

1.2.2 La modulation du Traffic Engineering aux contraintes des applications

Traditionnellement, le Traffic Engineering (TE) dans les réseaux est opéré indépendamment d'applications ou services particuliers. Néanmoins dans les travaux de Jiang et Al. [JZSRC09], les auteurs étudient une nouvelle approche de coordination qui formalise la gestion du Traffic Engineering (pour sélectionner les chemins réseau les plus efficaces pour le trafic) par rapport à la problématique de sélection de serveur (pour sélectionner le serveur le plus approprié pour chaque client). Ils ont défini un ensemble de modèles caractérisés par des niveaux de coopération différents entre les opérateurs réseaux et les fournisseurs de services : un modèle où chaque opérateur est indépendant dans sa gestion, un modèle de coopération intermédiaire où les acteurs s'échangent des informations et un dernier modèle où la coopération fusionnelle, c'est-à-dire que les deux acteurs ne font plus qu'un. Dans chaque modèle, les auteurs étudient les performances des techniques d'optimisation qu'ils proposent et analysent leurs utilités.

Les modèles du système Les modèles étudiés par les auteurs correspondent à trois situations différentes ; (1) aucune coordination entre réseau et service, (2) une coordination limitée avec systèmes de décision indépendants de part et d'autres et (3) une coordination entière avec systèmes de décision fusionnés. Ces scénarios s'appuient sur les interactions entre les systèmes TE et SS (Sélection de Serveurs) illustrée dans la figure 1.10.

L'approche d'optimisation L'intérêt des études qui ont été menées dans cet article réside dans la résolution conjointe des deux problématiques de TE et SS. Les auteurs ont spécifié le problème mathématique qui modélise chaque problématique. Ainsi le premier problème propose de minimiser le paramètre TE qui est la somme des coûts d'utilisation des liens du réseau. Ces coûts dépendent des taux d'utilisation des liens. Le second problème propose de minimiser le paramètre SS qui est la somme des délais expérimentés par les flux de service. Les auteurs s'appuient ensuite sur le diagramme de Pareto [Par] pour résoudre le problème global et analyser les interactions entre TE et SS.

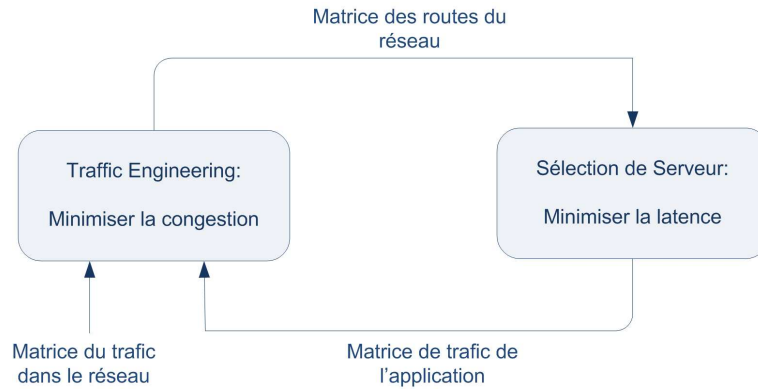


FIGURE 1.10 – Illustration des interactions entre les systèmes TE et SS

1.2.3 Synthèse des techniques étudiées

Les techniques que nous avons présentées dans les sections précédentes montrent qu'il n'y a pas de solution unique à la problématique que nous avons posée. La plupart des approches se basent sur une interface centralisée, avec ALTO souvent explicitement cité, pour la gestion du trafic en fonction d'informations d'état sur les réseaux de transport. Ces différents systèmes reposent aussi sur facteurs d'optimisation différents, comme la distance au sens réseau entre les noeuds applicatifs ou bien la charge de trafic sur les différents liens. Nous synthétisons ces différentes propositions dans le tableau récapitulatif 1.1.

1.3 Les interfaces issues du monde de la normalisation

Le besoin d'assurer une meilleure gestion des flux de service dans les réseaux a poussé les acteurs du monde des Télécommunications et plus récemment les géants du Web à travailler sur la spécification d'interfaces capables de fournir informations et fonctionnalités disponibles dans le réseau aux applications. Fabricants d'équipements, opérateurs de réseaux et autres acteurs ont joint leurs efforts dans la normalisation des ces interfaces pour assurer une pérennité et garantir une interopérabilité entre leurs systèmes. Nous aborderons dans les sections qui suivent les principaux systèmes, architectures et interfaces qui ont fait l'objet de travaux poussés dans les divers organismes de normalisation, ITU, ETSI, 3GPP, IETF, etc.

1.3.1 Le Next Generation Network

Le réseau de nouvelle génération, Next Generation Network (NGN) [TcSfANT09a], est un réseau à commutation de paquets capables de transporter les données de différents services y compris les services de télécommunications. Il se caractérise par sa capacité à fédérer différentes technologies de transport, à offrir des mécanismes de ges-

TABLE 1.1 – Synthèse de différentes propositions académiques pour la problématique de gestion trafic par interface entre services applicatifs et réseaux

Approche	Interface	Facteur d'optimisation	Technique de résolution	Applications
P4P	P4P (proche d'ALTO)	Trafic intra-domaine	Connections aux pairs les plus proches	P2P
ALTO	Protocole ALTO	Trafic intra et inter domaines	Non définie	P2P et CDN
PaDIS	ALTO	Distance, Latence, Bande passante	Ordonnancement des sources par priorité effectué par l'opérateur	P2P et CDN
CaTE	ALTO	Utilisation maximale des liens	Équilibrage des flux	CDN
DONAR	Non évoquée	Distance, Latence	Algorithme distribuée de redirection	CDN
NAPA-WINE	ALTO	Distance	Connections aux pairs les plus proches	P2P
I-CDN	Interne à l'opérateur	Charge de trafic	Équilibrage de la charge	CDN
Coordonnées Virtuelles	protocole propre à chaque technique	Distance	Redirection au noeud le plus proche	P2P et CDN
TE	Interface interne	utilisation des liens	Modulation du TE	CDN

tion de qualité de service et à garantir une indépendance entre services et technologies de transport.

L'organisme de normalisation Telecommunications and Internet converged Services and Protocols for Advanced Networking (TISPAN) [TIS12], rattaché à European Telecommunications Standards Institute (ETSI) est responsable des travaux de spécifications du NGN. La première version (*Release 1*) qui a été achevée en décembre 2005, a fourni les normes nécessaires au développement et essais de la première génération de réseaux NGN. Elle a permis d'adopter l'architecture de l'IP Multimedia Subsystem (IMS) [Her10, PMKN04].

La deuxième version (*Release 2*) a été achevée au début de 2008, et a ajouté des éléments clés pour les réseaux NGN comme le support du Internet Protocol television (IPTV) pour les systèmes IMS et non-IMS, l'extension des travaux aux réseaux domestiques et terminaux, ainsi que l'interconnexion avec les réseaux d'entreprises.

1.3.1.1 Architecture fonctionnelle du NGN

L'architecture fonctionnelle du NGN [TcSfANT09a] exposée dans la figure 1.11 est structurée en deux couches ; une couche de service et une couche de transport. La couche de service comprend les composants suivants :

- le noyau du sous-système multimédia IP (IMS)
- le sous-système d'émulation du PSTN/RNIS
- les autres sous-systèmes multimédias (IPTV...)
- les composants communs utilisés par les différents sous-systèmes tels que ceux requis pour la gestion des profils utilisateurs ou pour les opérations de comptage.

Cette architecture particulière a été pensée dans l'expectative de pouvoir intégrer et inter-opérer avec de nouveaux sous-systèmes couvrant de nouvelles exigences et types de service.

1.3.1.2 Le contrôle du transport dans le NGN

La couche de transport de l'architecture présentée dans la figure 1.11 comprend une sous-couche de contrôle du transport au-dessus des fonctions de gestion du transport dans les réseaux accès et cœur. Cette sous-couche de contrôle est composée de deux sous-systèmes :

- Le sous-système de connexion au réseau, Network Attachment SubSystem (NASS)
- Le sous-système de contrôle d'admission et ressources, Resource and Admission Control Subsystem (RACS)

Le sous-système NASS Le NASS fournit certaines fonctionnalités au niveau accès comme l'adressage IP dynamique, la configuration des terminaux, l'authentification et autres autorisations liées, etc. Le NASS dispose d'interfaces avec d'autres modules fonctionnels (voir figure 1.12) :

- L'équipement utilisateur (interface e3) pour l'accès et la configuration

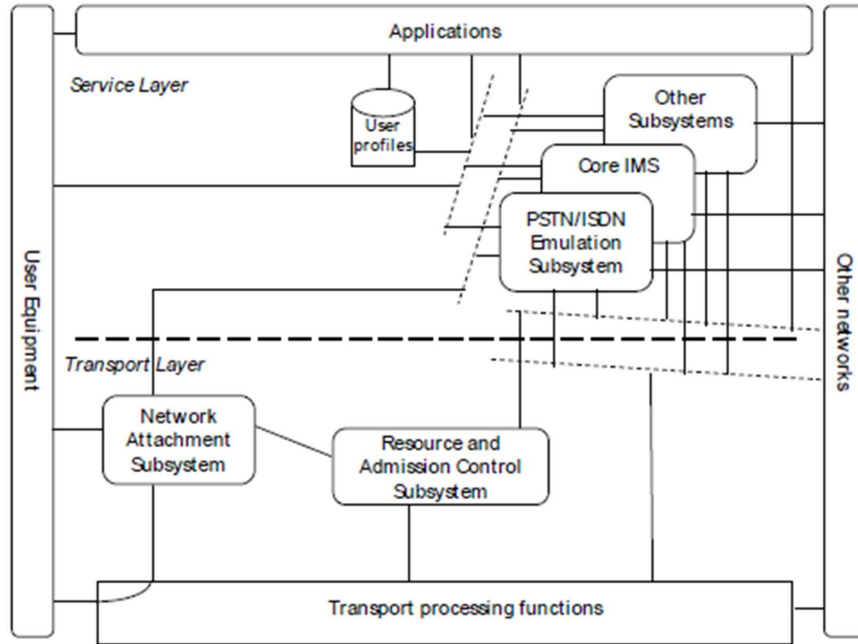


FIGURE 1.11 – L'architecture fonctionnelle du système NGN

- Les fonctions de gestion de transport (interfaces a1 et a3) pour les besoins de contrôle d'adressage, d'authentification, autorisation ,etc.
- Le RACS (interface e4) pour exporter les informations liées aux profils et privilèges des utilisateurs
- Les sous-systèmes de contrôle implémentés par les services pour délivrer les informations et notifications relatives aux sessions des utilisateurs

Le document [TcSfANT09b] fournit des informations plus détaillées sur les spécifications et les interfaces du NASS.

Le sous-système RACS RACS est le sous-système NGN responsable du contrôle des mécanismes d'applications des politiques et règles de réservation de ressources et d'admission que ce soit pour le trafic unicast ou multicast dans les réseaux d'accès ou coeur. En plus d'assurer ces fonctions, le RACS prend également en charge le contrôle des translations d'adresses et port réseaux, "*Network Address and Port Translation*" (NAPT). Il est capable aussi de la gestion des aspects liés à l'établissement et la modification des politiques d'acheminement du trafic, de la qualité de service de bout en bout et des besoins de comptage. Le RACS dispose d'interfaces (voir figure 1.12) avec :

- Le NASS (interface e4) pour récupérer les informations liées aux profils des utilisateurs
- Les sous-systèmes de contrôle implémentés par les services (interface Gq') pour

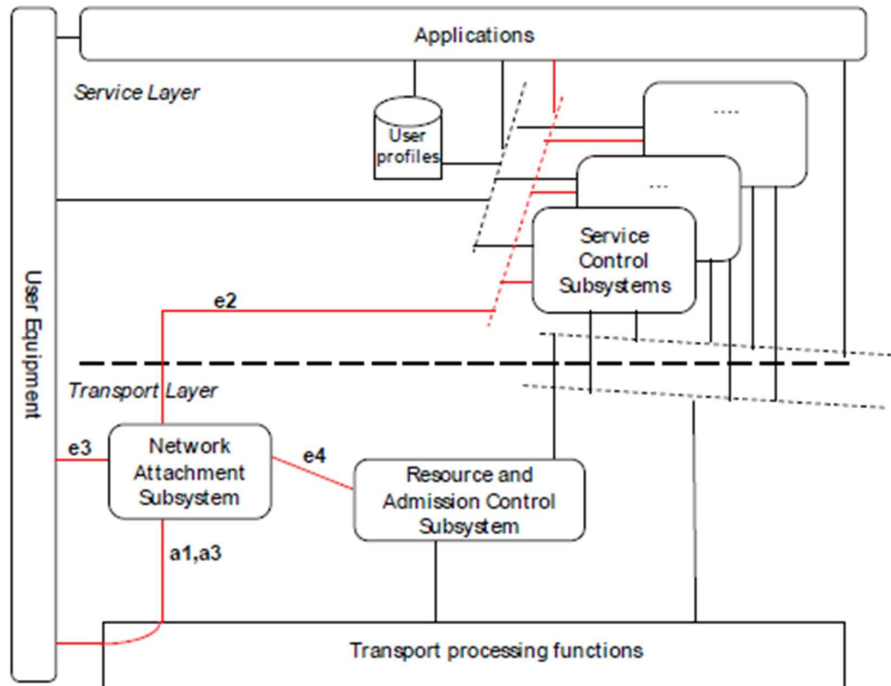


FIGURE 1.12 – Illustration des interfaces externes du NASS

exposer les différents services qu'il fournit

- Les fonctions de gestion du transport (interfaces Re et Ia) pour l'application et contrôle des procédures de gestion du trafic

Le document [TcSfANT10] fournit des informations plus détaillées sur les spécifications et les interfaces du RACS.

La gestion du transport dans le NGN Divers fonctions sont responsables de la gestion du transport dans les réseaux d'accès et de coeur pour le NGN. Elles implémentent à la fois des mécanismes élémentaires comme la transmission, acheminement ou routage des paquets et d'autres fonctions définies comme des entités fonctionnelles. On cite les fonctions suivantes :

- *Media Gateway Function* (MGF), fournit les fonctionnalités de traitement des sessions multimédias comme le transcodage.
- *Border Gateway Function* (BGF), assure une interface entre deux domaines de transport IP et fournit des fonctionnalités comme la translation d'adresse ou de port (NAPT).
- *Resource Control Enforcement Function* (RCEF) pour l'application des règles d'acheminement des flux

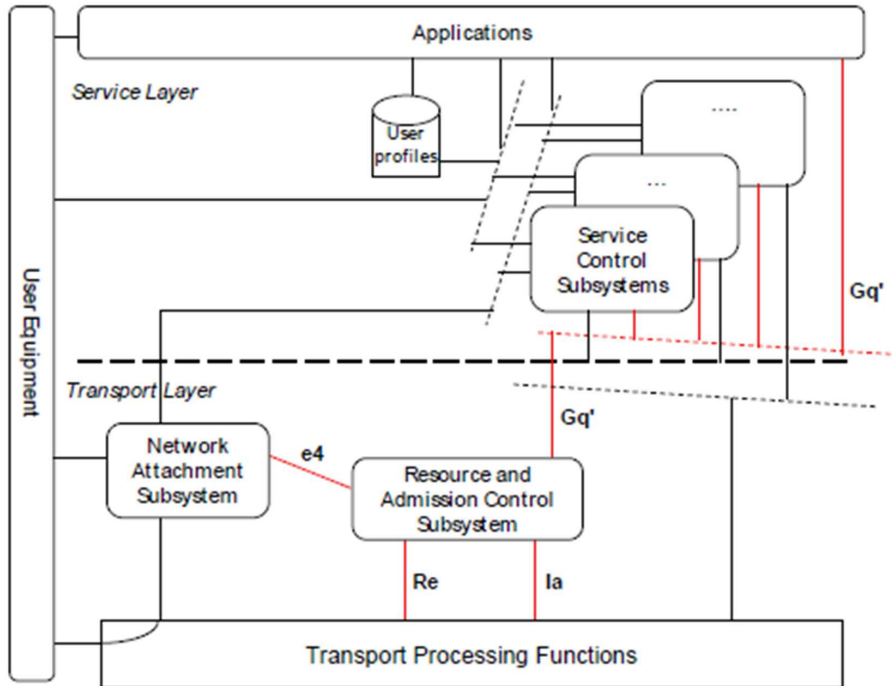


FIGURE 1.13 – Illustration des interfaces externes du RACS

- *Access Relay Function* (ARF), agit comme un relai entre l'équipement utilisateur et le NASS pour transmettre les requêtes de connexions au réseau par exemple.
- *Signalling Gateway Function* (SGF), assure la gestion de la signalisation au niveau transport comme la conversion ou adaptation entre domaines distincts.
- *Multimedia Resource Function Processor* (MRFP), fournit des fonctionnalités plus complexes que la MGF telles que les services de conférence, analyse de contenus, etc.
- *Access Management Function* (AMF), convertit les requêtes de connexion émises par l'équipement utilisateur en un format compréhensible par le NASS
- *Basic Transport Function* (BTF), fournit des fonctionnalités de transmission et contrôle des paquets.

Le sous-système RCEF Le sous-système RCEF est une entité de gestion de transport qui implémente une ou plusieurs des fonctions élémentaires suivantes :

- Le filtrage des paquets en fonction des adresses IP ou ports pour autoriser uniquement le trafic ayant droit
- Le marquage des paquets pour le trafic sortant
- L'admission du trafic entrant
- La répartition de ressources entre les flux de trafic pour s'assurer qu'ils restent dans les limites autorisées

Le RCEF, en tant qu'entité fonctionnelle, peut être implémentée aux niveaux des différents niveaux du réseau à savoir l'accès, l'agrégation, ou le coeur. Plusieurs instances du RCEF peuvent cohabiter dans le même domaine. Le RACS interagit avec le RCEF à travers les interfaces *Ia* et *Re* pour commander les différentes fonctions élémentaires fournies par ce dernier.

1.3.2 L'IP Multimedia Subsystem

1.3.2.1 Présentation de l'IMS

L'IMS, "*IP Multimedia Subsystem*" [Her10, PMKN04], est un système conçu pour fournir un ensemble de services plus riches comme la VoIP, Video-Conférence, etc, avec une plate-forme standardisée. Initialement proposé par le 3GPP, "*3rd Generation Partnership Project*" [3GP12] pour les réseaux sans fil, l'IMS vise à assurer la convergence entre les différents réseaux d'accès (convergence entre réseaux fixes et mobiles) en offrant les mêmes services quel que soit le type de réseau d'accès auquel le terminal est connecté, sur la base du protocole IP. L'IMS devrait également permettre la mobilité des utilisateurs.

1.3.2.2 L'architecture 3GPP PCC

L'architecture 3GPP Policies and Charging Control (PCC) [rGPP12] a été conçue afin de remplir deux fonctions principales :

- Le comptage au niveau flux de données avec un contrôle des procédures du comptage et de facturation et de gestion du crédit en direct.
- Le contrôle des politiques de gestion de flux comme les politiques d'admissions et de gestion de qualité de service (QoS).

L'architecture 3GPP PCC spécifie les modules fonctionnels suivants : le "*Policy and Charging Enforcement Function*" (PCEF), le "*Bearer Binding and Event Reporting Function*" (BBERF), le "*Policy and Charging Rules Function*" (PCRF), le "*Application Function*" (AF), le "*Traffic Detection Function*" (TDF), le "*Online Charging System*" (OCS), le "*Offline Charging System*" (OFCS) et le "*Subscription Profile Repository*" (SPR) ou autrement appelé le "*User Data Repository*" (UDR).

Le sous-système AF Le module AF permet d'inter-connecter l'architecture PCC associée au réseau à commutations de paquets à la couche service du système IMS. Il assure certaines fonctions comme la corrélation des identifiants de comptage entre les systèmes IMS et PCC, la notification des événements sur les sessions (terminaison, modification) au PCRF, la réception de notifications concernant les canaux de transport gérés par le PCRF, etc.

Le sous-système PCRF Le PCRF est responsable principalement de la réservation des ressources et du contrôle d'admission pour le transport des flux de données dans le réseau mais il peut assurer aussi des fonctions de NAT. En mode *push*, le PCRF

reçoit des demandes de réservation de ressources de l'IMS, transmises par le module AF. Le PCRF se charge aussi de vérifier si les demandes de ressources correspondent bien aux privilèges accordés à l'utilisateur. Il accède à cet effet à la base du système HSS de l'IMS. En outre, il coordonne les demandes d'admission transmises aux PCEF et BBERF. Le PCRF et l'architecture globale 3GPP pour le comptage et les politiques de contrôle sont détaillés dans le document 3GPP TS 23.203 [rGPP12].

Le sous-système PCEF Le PCEF est responsable de l'application des demandes de réservation de ressources ou du traitement des paquets et flux de données (filtrage, marquage, modification, etc), transmises par le PCRF à travers l'interface de *Gx*. Les règles PCC peuvent être pré-définies dans le PCEF et en conséquence référencées par le PCRF à travers l'interface *Gx*. Elles peuvent aussi être définie de façon dynamique. Une règle PCC comprend un certains nombres d'identifiants, de paramètres et de méthodes relatives aux flux et à leur transport dans le réseau.

Le sous-système BBERF Le BBERF offre les mêmes fonctionnalités que le PCEF tout en assurant la capacité d'associer les flux de données aux canaux ou porteuses du système radio. Le BBERF est piloté par le PCRF à travers l'interface *Gxx*.

Deux mécanismes de fonctionnement sont définis pour le PCC :

- Un mode *pull*, indépendant de l'IMS, dans lequel le BBERF et PCEF transmettent aux PCRF les politiques d'établissement des canaux de transport et la configuration des modes de transmission et de comptage. Ce mécanisme permet par exemple le contrôle des paramètres QoS dans les réseaux "Evolved Packet Systems" (EPS) pour tout service IP.
- Un mode *push*, dans lequel un module IMS (par exemple le module AF), définit les politiques de comptage et de QoS transmises au PCRF lors de l'établissement de la session IMS. Le module AF, défini dans le cadre de l'EPS, n'est pas une entité autonome mais il est habituellement implémenté dans un P-CSCF (à ne pas confondre avec le module "Application Server" (AS) défini dans l'IMS).

La gestion des flux par le PCC pour l'IMS Le PCC offre divers fonctionnalités pour la gestion des flux de données pour les services IMS :

- L'association des flux de données aux canaux de transport (par exemple porteuses radio) des systèmes IP-CAN ("*Connectivity Access Network*").
- Le filtrage des flux, à savoir bloquer ou autoriser la transmission les paquets appartenant à un flux de données ou identifiés par un identifiant de service.
- La notification d'événements concernant les données et les ressources dans les équipements relais comme le PCEF.
- La gestion de la QoS associée aux flux et services, à savoir la spécification de la QoS dont bénéficie un flux de données, une application particulière ou des canaux de transports des systèmes IP-CAN.

- La redirection des paquets appartenant à une application particulière à l'adresse de redirection spécifiée.
- L'initiation des canaux de transport pour les systèmes IP-CAN qui supportent une initiation du réseau de leurs canaux.

1.3.3 L'architecture commune 3GPP et TISPAN pour l'intégration de l'IMS et le NGN

L'architecture EPS du 3GPP met particulièrement l'accent sur les réseaux mobiles et sans fil. Dans ce contexte, le rôle de l'organisme TISPAN fût de se concentrer plutôt sur les réseaux fixes. L'architecture commune qui a été définie par le 3GPP et TISPAN comme base à leurs travaux est décrite dans le document [TcSfANT09a]. Cette architecture, illustrée dans la figure 1.14, définit deux couches fonctionnelles, chacune comprenant un certain nombre de sous-systèmes :

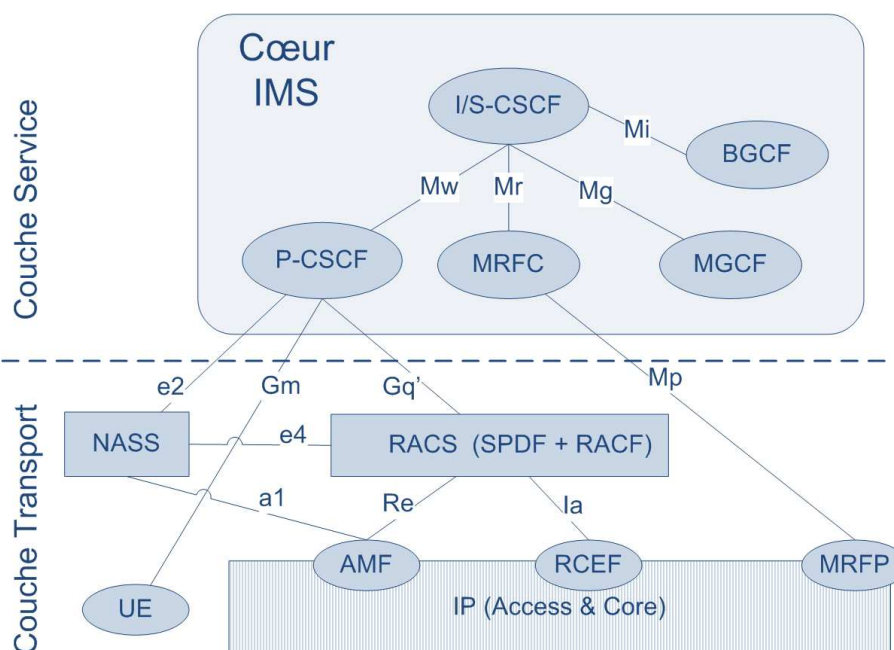


FIGURE 1.14 – Illustration de l'architecture commune 3GPP/TISPAN

- La couche service, comprenant le cœur IMS (CSCF, MRFC, etc). Elle fournit toutes les fonctionnalités nécessaires au fonctionnement des services IMS.
- La couche de transport, intégrant le sous-système NASS, qui contrôle l'accès au réseau IP à travers les différentes couches de transport physiques et le sous-système RACS qui contrôle les politiques d'admissions et d'utilisation des ressources du réseau.

Les modules et interfaces TISPAN ne sont pas entièrement alignés sur celles du 3GPP. La tendance des organismes de normalisation est d'aligner progressivement leurs

systèmes à l'exemple du RACS qui apparaîtra comme une instance spécifique du PCRF défini par le 3GPP.

1.3.4 OSA/Parlay

1.3.4.1 Introduction à OSA/Parlay

Traditionnellement, l'écosystème des télécommunications est un environnement assez cloisonné. De ce fait, les services et applications développés pour ce genre d'écosystème ne peuvent être développés qu'en interne grâce à la maîtrise des différentes technologies réseau déployées. Des efforts ont été investis dans la perspective d'ouvrir cet environnement en offrant la possibilité de considérer certains aspects liés aux réseaux dans le développement d'applications [MK03]. De ces initiatives d'ouverture ont émergé de nouveaux modèles économiques permettant à des entreprises tierces de développer et fournir des services, exclusifs jusqu'alors aux opérateurs réseaux. Ces services peuvent bénéficier de fonctions enrichies par des fonctionnalités du réseau par le biais d'interfaces de programmation, Application Programming Interface (API), avec des outils et technologies tels que Java ou Web Services.

Les APIs de programmation Parlay [LLOH05, UVB06], autrement connu sous "*Open Service Access*" (OSA), est un ensemble d'interfaces ouvertes et standardisées qui permettent aux applications d'accéder à certaines ressources du réseau. L'approche Parlay permet d'interfacer les fonctionnalités des réseaux de télécommunications avec les applications et services via des interfaces sécurisées, efficaces et offrant des capacités de comptage et facturation. Les APIs OSA/Parlay ne dépendent pas d'un réseau particulier mais ont été conçues dans la perspective de permettre aux applications d'être hébergées par les opérateurs réseaux et par les fournisseurs de services tiers.

Les spécifications des APIs OSA/Parlay ont été menées conjointement par l'ETSI [ETS02], le 3GPP [rGPP09] et le groupe Parlay. Les spécifications se sont basées sur un modèle UML, à partir duquel le code en java des APIs a été développé. En utilisant cette approche, les travaux réalisés par l'ETSI, Parlay, et le 3GPP ont été homogènes.

Des APIs en Web-services, connues sous le nom *ParlayX* sont également disponibles. Elles ont été normalisées conjointement par l'ETSI, Parlay et le 3GPP aussi. ParlayX est une interface Web Service permettant l'accès aux fonctionnalités et ressources des réseaux de communications.

1.3.4.2 Présentation des APIs

Les APIs OSA/Parlay fournissent une couche d'abstraction qui ne dépendent pas d'une technologie particulière et qui offre des fonctionnalités telles que la gestion d'appels, la localisation ou l'interaction avec les utilisateurs. Les APIs ParlayX offrent également les mêmes fonctionnalités, mais stimulent davantage l'intérêt des développeurs d'applications qui n'ont pas des connaissances poussées des technologies télécom. En effet, ParlayX offre un niveau d'abstraction meilleur par rapport aux APIs OSA/Parlay, en profitant d'interfaces de type Web-Services plus répandues.

L'usage principal des APIs Parlay a concerné dès le départ les services de téléphonie, qui nécessitent des fonctionnalités de gestion et contrôle d'appels, tant pour les réseaux fixes [JBA05] que mobiles [JG05]. Des fonctionnalités avancées du système NGN incluant la gestion de la QoS, le multicast, la réservation de ressources (notamment pour permettre à des services de type TV de réserver une bande passante à la demande) ont été ouvertes aux applications à travers les APIs,.

1.3.4.3 OSA/Parlay Connectivity Manager

L'API "*OSA/Parlay Connectivity Manager*" est l'interface qui permet aux services d'accéder aux fonctionnalités de QoS dans un réseau. Les APIs OSA/Parlay plus globalement définissent un cadre de fonctionnement (le *framework*) et un certain nombre de fonctionnalités de services appelées "*Service Capabilities Features*" (SCFs). Le framework est le point de contact initial pour les applications. Il leur permet de découvrir et accéder aux fonctionnalités offertes par le réseau. Ce framework dispose de mécanismes appropriés pour assurer la sécurité des interfaces à travers des fonctions d'authentification, autorisation, gestion des privilèges, etc.

Les SCFs fournissent l'ensemble des interfaces qui permettent aux applications d'accéder aux fonctions offertes par le réseau. L'API *Connectivity Manager* définit un ensemble de SCFs qui permettent une gestion conjointe des paramètres de QoS par les fournisseurs de services et opérateurs réseaux qui les utilisent. Les mécanismes de contrôle de QoS dans le réseau ne sont pas imposés mais les techniques de différenciation de qualité de services de type "Differentiated Services Code Point" (DSCP) ont été évoquées pour la gestion d'agrégat de flux.

Pour assurer cette gestion de QoS dans le réseau, l'opérateur réseau maintient une liste des sites des fournisseurs de services avec leurs "*Service Access Points*" (SAP) pour distinguer leurs flux de données. L'objectif est de permettre aux fournisseurs de services de définir la configuration QoS pour leurs flux de données entre les SAPs. Ces flux sont appelés "*Virtual Provisioned Pipes*" (VPrP). Un groupe de VPrPs est appelé "*Virtual Provisioned Network*" (VPrN). Les applications d'un fournisseur de services avec des VPrNs établis utilisent par la suite l'API pour construire des VPrPs pour leurs flux de données avec des paramètres QoS précis.

Le sous-système RACS du NGN, responsable de la gestion des mécanismes de contrôle d'admission, de réservation de ressources, de gestion de la QoS dans le réseau, de l'application des politiques de filtrage, etc, constitue la cible des travaux de l'ETSI [ETS02] en vue de l'intégration de l'API *OSA/Parlay Connectivity Manager* pour l'ouverture des réseaux NGNs aux services tiers et la mise à disposition des divers fonctionnalités qu'ils implémentent.

1.3.5 GSMA OneAPI

1.3.5.1 Introduction à OneAPI

L'initiative OneAPI du GSMA [GSM12] a abouti à la spécification d'un ensemble d'APIs très orientée Web pour permettre aux opérateurs de réseaux mobiles ou autres d'exposer des informations utiles et offrir des fonctionnalités réseau aux développeurs d'applications (voir figure 1.15 [Unm11]).

Pour la spécification de OneAPI, le GSMA a collaboré avec le "*Open Mobile Alliance*" (OMA) [OMA12] afin de revoir et ré-utiliser des APIs existantes telles que les APIs ParlayX orientées Web-Services. Les efforts conjoints des deux groupes ont permis de revoir les spécifications de ces APIs et les étoffer pour offrir une utilisation beaucoup plus facile aux développeurs d'applications. A titre d'exemple, parmi les voies d'amélioration empruntées, on peut citer l'utilisation de l'approche RESTful et des schémas de présentation JSON et XML, très connus des développeurs d'applications web.

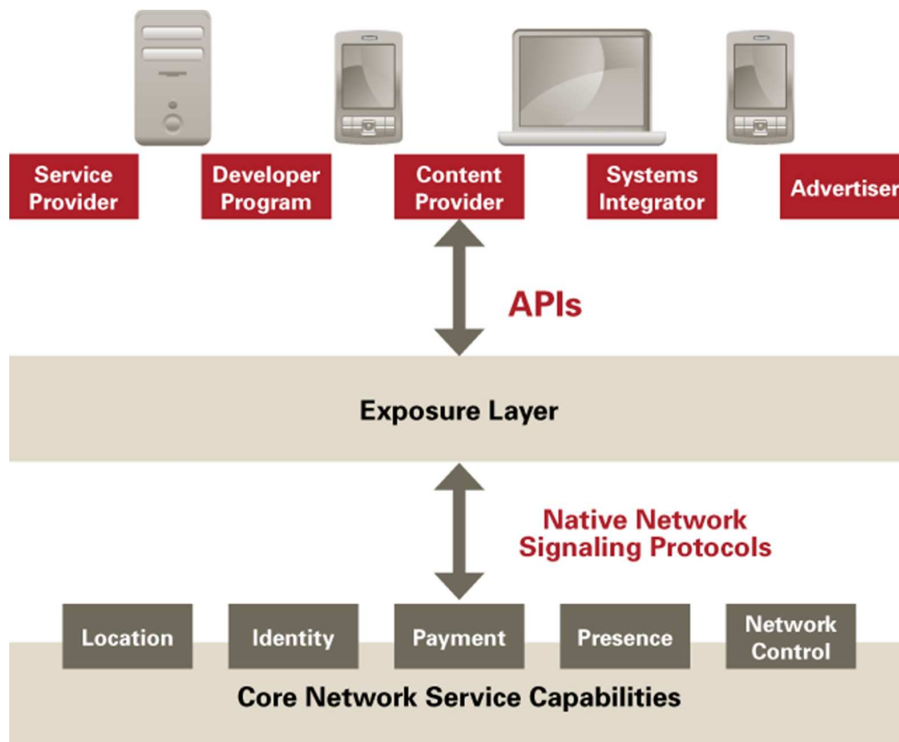


FIGURE 1.15 – Illustration des interfaces produites par les APIs OMA entre services et réseaux

1.3.5.2 Les APIs OneAPI

Cette collaboration a abouti à la standardisation entière de OneAPI [Ope12] comme profil d'APIs dans le standard *REST NetAPIs* du OMA. OneAPI se compose d'un ensemble d'APIs qui permettent l'accès à certaines informations et ressources du réseau

(APIs de *messaging*, de localisation, de paiement, etc). Ces APIs ont été conçues de manière à offrir une intégration facile pour les opérateurs réseaux d'un côté et une interface unifiée et simple aux développeurs d'applications d'un autre côté. Les opérateurs réseaux sont en effet en mesure d'implémenter OneAPI sans apporter d'énormes modifications dans leurs architectures et systèmes. Les fournisseurs de services de leur côté auront un accès efficace aux informations et ressources relatives aux réseaux et utilisateurs à travers les différentes APIs.

L'API SMS par exemple permet à une application d'envoyer et de recevoir des messages en texte. Des méthodes supplémentaires ont été définies afin d'offrir des fonctionnalités avancées. Elles permettent aux applications par exemple de vérifier la bonne remise d'un message. Similaire à l'API SMS, l'API MMS permet d'envoyer et de recevoir des messages multimédia. Elle offre aussi des méthodes capables de demander l'acquittement de la bonne remise des messages, éventuellement à destination d'autres applications.

L'API de localisation quant à elle permet à une application de déterminer la localisation d'un ou de plusieurs terminaux mobiles. L'application utilise un certain nombre d'informations sur le profil de l'utilisateur pour permettre à l'opérateur réseau de l'identifier et le localiser. Ce dernier renvoie une réponse indiquant la position du terminal avec des données portant sur l'altitude, la latitude, la longitude et l'heure de localisation.

OneAPI met à disposition des applications aussi une API pour la facturation et le paiement des services ou contenus qu'elles offrent. Différentes techniques de paiement sont proposées.

Dans ses versions 2.0 et 3.0, OneAPI offre plus d'APIs mettant à disposition des développeurs et fournisseurs de services plus de fonctionnalités du réseau. Citons à titre d'exemple la possibilité de demander une QoS différenciée pour des services de diffusion vidéos.

1.3.6 Openflow

Openflow est une nouvelle conception des systèmes réseaux qui fait partie des travaux sur les "*Software Defined Networking*" (SDN) [MAB⁺08]. Ce standard ouvert sépare désormais le plan de données du plan de contrôle. En conséquence, le traitement des données est toujours effectué dans les équipements comme les *switchs* alors que le plan de contrôle est porté dans un autre équipement comme un serveur centralisé dans le réseau. Les actions de traitement des flux de données comme la redirection sur un port particulier ou la suppression de paquets sont prises par le plan de contrôle et communiquées au plan de données à travers un protocole spécifique comme le protocole Openflow.

Openflow offre ainsi l'avantage de pouvoir mieux maîtriser une infrastructure réseau distribuée grâce à un contrôle à distance standardisé. Il offre des perspectives pour la virtualisation des réseaux ou pour l'amélioration des techniques de sécurité. Cette approche agit par intervention sur les flux de données dans le réseau. Pour des services

internet, cela nécessite une ouverture des réseaux pour permettre un contrôle dicté par des tiers.

1.4 L'utilité d'une interface de collaboration

L'évolution remarquable des services sur Internet a été soutenue par un développement des technologies réseaux de dernière génération comme le LTE ou la fibre optique, la diversité des terminaux connectés (smartphones, tablettes, TV numérique, etc) et l'attractivité des offres de services dont nous citerons à titre d'exemple YouTube et Netflix. Dans ce contexte bien implanté, nous débattons des lacunes des approches existantes dans la gestion de la croissance du trafic internet et nous introduisons l'utilité de notre contribution.

1.4.1 Des interfaces inappropriées

La standardisation d'interfaces capables de faire le lien entre services et réseaux a mobilisé divers organismes de standardisation depuis le début des travaux sur les réseaux NGN et l'IMS. Ce cadre bien particulier de systèmes cloisonnés qui sont sous le contrôle total de l'opérateur de communications est incompatible avec le modèle ouvert des services sur Internet. En partant de ce constat, ces interfaces, entre RACS, NASS et autres PCC qui contrôlent le réseau, et les modules de services IMS ou autres, ont été définies pour les services qui seront déployés par l'opérateur. Il en résulte des obstacles considérables et de natures différentes quant à leurs utilisation aujourd'hui par des fournisseurs de services sur Internet. On cite par exemple les contraintes sécuritaires et financières. En effet, ces interfaces offrent des fonctionnalités de contrôle dans le réseau (contrôle d'admission, gestion de la QoS, filtrage de flux, etc). Une utilisation en interne ne pose pas de problèmes aux opérateurs mais la mise à disposition de certaines fonctionnalités aux tiers est truffée de risques. Financièrement, les modèles économiques des opérateurs télécom et ceux des fournisseurs de services sur internet sont différents. La convergence des intérêts autour de l'usage des interfaces réseaux-services n'est pas une équation facile à résoudre. Cependant, l'obstacle majeur qui réside dans cette voie est l'ensemble cadre et besoins techniques qui ont induit la spécification de ce type d'interfaces. On parle ici des fonctionnalités disponibles aux services des opérateurs télécom. Une éventuelle transposition aux services interne n'a pas été développée.

1.4.1.1 Les limites de la maîtrise de la QoS dans le réseau

Les principales fonctionnalités qui sont offertes par les interfaces définies par TIS-PAN pour le NGN [TcSfANT09a] et le 3GPP pour l'IMS [rGPP12] concentrent des services de gestion du comptage ou de la facturation et des services de maîtrise de la QoS dans le réseau. L'exploitation de ces fonctionnalités nécessite l'intervention de plusieurs modules impliquant l'utilisateur, le réseau et le service engageant de fait l'utilisation de ressources non négligeables. Cette logique de fonctionnement a été pensée pour assurer une QoS à des services majoritairement conversationnels et contrôlés par

l'opérateur. Le contraste est bien évidemment important quand il s'agit de services Internet. Les profils, les besoins et les contraintes sont différentes. En effet, apporter des garanties de QoS à des services sous le contrôle de différents fournisseurs tiers, dont les flux sont susceptibles de traverser plusieurs domaines réseaux, dont les modèles économiques sont différents, dont le développement et le déploiement répond à des règles différentes de ceux des opérateurs, ne paraît pas adapté. Par ailleurs, il est tout aussi important de noter que ces garanties de QoS n'apportent pas d'éléments d'améliorations à la problématique de croissance de trafic dans les réseaux à tous les niveaux. On peut s'attendre à vrai dire à l'effet inverse car la généralisation de l'utilisation de fonctionnalités de gestion de QoS dans le réseau pour les services Internet nécessitera la mobilisation d'énormes ressources, et ce non pas pour faire face à la croissance du trafic mais pour satisfaire les garanties QoS de flux auparavant transportés en mode "*best effort*". Nous partons de ce constat pour attirer la lumière sur le fait que la croissance du trafic est un handicap qui affectera à terme les performances à la fois des réseaux et des services qui les utilisent. Une gestion efficace de ce trafic nécessite une collaboration étroite entre les acteurs concernés. Cependant, cette collaboration doit s'inscrire dans un cadre approprié qui satisfait les opérateurs de réseaux et les fournisseurs de services.

En ce sens, l'initiative de OneAPI [Ope12] qui a pris la relève du OSA/Parlay [ETS02, rGPP09] a cherché depuis le début à concilier l'environnement des applications web et internet avec les contraintes des opérateurs de communications à travers la spécifications d'interfaces de programmation APIs qui font l'usage de technologies à portée des fournisseurs de services comme les *Web-Services*, *JSON*, *XML* ou encore l'approche *RESTful*. Cette démarche a facilité l'implantation d'une interface entre réseaux et services afin de permettre à ces derniers de bénéficier d'un certain nombre de fonctionnalités et/ou services disponibles dans le réseau. Malheureusement ces APIs sont portées sur l'aspect service avec la possibilité de bénéficier des fonctionnalités de comptage ou de messagerie. Les travaux de la troisième génération de OneAPI font état de fonctionnalité de réservation de ressources dans le réseau pour certains services multimédias. Ces travaux ne portent aucune attention particulière sur la croissance du trafic des services susceptibles de bénéficier des APIs OneAPI.

1.4.1.2 Un intérêt convergent vers la collaboration entre réseaux et services dans la littérature

Les travaux amorcés par le projet P4P [XKSY, XYK⁺08] ont tracé une nouvelle direction dans les possibilités d'amélioration de la gestion du trafic et de la qualité des services sur Internet. A l'IETF, le groupe de travail ALTO [ALT12] a été créé afin de normaliser une interface capable de garantir un échange d'informations fructueux entre le réseau et les applications qui l'utilisent. L'objectif est de réaliser un acheminement du trafic au niveau applicatif qui soit adapté aux informations exposées au préalable par le réseau. En effet, divers travaux qui ont précédé cette proposition se sont basés sur des systèmes de coordonnées virtuelles. Ces systèmes se basent principalement sur des estimations de latence à travers les *RTT* pour segmenter les domaines réseaux et

localiser les entités applicatives. Outre notre critique envers l'exactitude des résultats dérivés de ces estimations, ces travaux ne peuvent égaler les performances d'un système qui puise ses informations à la source, c-à-d chez l'opérateur qui contrôle le réseau.

Très récemment, d'autres travaux [WJFR10, FPS⁺12, WLCM11, KMK⁺10, PFA⁺10] ont développé cette dimension de collaboration entre réseaux et applications. L'architecture adoptée dans certains de ces travaux est explicitement ou implicitement basée sur le protocole ALTO. Ils héritent en conséquence de deux inconvénients majeurs inhérents à ALTO. Le premier inconvénient provient de l'aspect unidirectionnel de l'interface. Il s'agit d'une exposition d'informations depuis le réseau vers le service applicatif. Cet aspect peut entraver une personnalisation de ces informations suivant les contraintes des applications. Le deuxième inconvénient est lié au premier car cet aspect unidirectionnel empêche aussi toute tentative d'optimisation par le réseau car tout simplement ne disposant pas d'informations. Au final, le réseau est incapable de mutualiser ses actions ou ses informations par rapport aux différents services applicatifs avec lesquels ils collabore dans l'optimisation du trafic et l'amélioration de leur qualités.

1.4.2 L'utilité d'un système de collaboration entre réseaux et services

Les nouvelles applications multimédias deviennent au fil du temps de plus en plus complexes et complètes avec une meilleure interactivité, une qualité en hausse grâce à des présentations *HD*, *3D* et bientôt *4K*, des fonctionnalités avancées, etc. Les infrastructures de distribution de contenus sont amenées à supporter une charge croissante et deviennent en conséquence de plus en plus distribuée. Dans ce contexte, les applications P2P offrent aussi des avantages indéniables et proposent des alternatives pour le partage de contenus. Ces tendances aboutissent au final à des besoins en transport et en qualité sans précédent sur les réseaux sous-jacents pour assurer une exécution de qualité des services engagés entre des entités applicatives imprévisibles.

En conséquence, au lieu de traiter le problème par un investissement colossal en surdimensionnant les ressources de transport, nous préconisons d'adopter des techniques intelligentes basées sur une collaboration entre les réseaux et les applications qui les utilisent afin d'employer de façon optimale les ressources et les services disponibles dans le réseau en veillant aussi à ce que les applications s'adaptent aux différentes contraintes. Elles bénéficieraient en contre-partie d'une QoE améliorée grâce à un transport efficace des données entre leurs entités. Deux applications phares sont particulièrement étudiées dans nos travaux, les CDNs et les P2P. Ces applications concentrent en effet aujourd'hui la majorité du trafic multimédia et présentent des opportunités d'optimisation concrètes.

1.4.2.1 L'intérêt pour les services CDNs

Les réseaux de distribution de contenus CDNs maintiennent plusieurs points de présence géographique pour rapprocher les serveurs de distributions aux utilisateurs. Ces serveurs peuvent ne pas être connectés directement aux réseaux de transport mais interconnectés comme une entité indépendante ou à travers d'autres réseaux. Ils peuvent

aussi être rattachés directement à certains noeuds des réseaux de transport, généralement au niveau des points de présence du réseau ou bien au niveau des noeuds multi-médias. Les contenus demandés par les utilisateurs sont répliqués dans ces serveurs de distribution pour offrir un accès plus rapide et plus fiable pour les clients. En effet, d'un côté le temps d'accès aux informations se trouve amélioré et d'un autre côté on évite un acheminement des informations depuis le serveur d'origine vers des clients lointains ce qui se traduit par une économie non négligeable d'utilisation de ressources réseaux. Les politiques de déploiement des contenus dans les divers noeuds de distribution appartenant aux CDNs peuvent varier en fonction de différents critères. Il existe des approches qui tentent d'anticiper les demandes des clients, d'autres qui se contentent de recevoir une requête explicite et d'autres approches qui cherchent à optimiser la distribution des contenus sur les différents serveurs en fonction de critères spécifiques comme la charge des noeuds de distribution par exemple. Nous nous concentrons dans nos travaux sur l'optimisation du transport du trafic réseau et l'amélioration de la qualité de service étant donné un contexte réseau et service connu.

Dans ce contexte de service, les opérateurs CDNs ne disposent pas aujourd'hui d'informations sur le contexte des réseaux de transport qui vont acheminer leur trafic vers les clients. Typiquement, un opérateur CDN dispose d'un très grand nombre de noeuds de distribution. Akamai annonçait en 2012 contrôler plus de 73,000 noeuds [CRM⁺10]. Même en réduisant ce nombre à une liste réduite de noeuds grâce à des informations de localisation qui permettent d'approcher la situation géographique d'un client, les opérateurs restent confrontés à une diversité de choix au moment de rediriger le client vers un noeud particulier. Il s'agit de faire un choix efficace qui permet d'optimiser un ou plusieurs aspects comme le transport des données entre le client et son noeud de service. Divers critères peuvent donc intervenir dans ce processus décisionnel. Certaines approches peuvent se baser par exemple sur le coût économique de transiter les données par tel ou tel réseau, le type de contenu, le type de terminal, etc. Il est difficile donc aujourd'hui d'affirmer ou d'infirmer une diversité de choix qui pourrait se présenter aux opérateurs CDNs lors de la sélection d'un noeud de distribution pour un client. Cet aspect dépend d'un ensemble de facteurs. Néanmoins dans les études qu'ils ont menées sur le trafic réel de deux fournisseurs d'accès à internet [PFA⁺10], les auteurs ont établi l'existence d'une diversité de noeuds de distribution capables de servir les clients d'un certain nombre d'acteurs majeurs dans les CDNs. En conséquence, une diversité de chemins réseaux s'offrent à eux.

Prenant en considération ces constatations et sans connaissance éclairée des conditions de transport dans les réseaux qu'ils utilisent pour l'acheminement de leur trafic, les opérateurs CDNs sont incapables de fournir un processus de décision fiable et optimal pour le choix du serveur le plus appropriée compte tenu de l'état des différents chemins de bout en bout entre le client et les serveurs disponibles. Certains opérateurs CDN [NSS10] tentent de remédier à ce problème en effectuant des mesures basées sur des techniques telles que le calcul du temps aller-retour dans les réseaux entre certaines entités applicatives qu'il s'agisse des clients eux mêmes, des noeuds de distribution

ou des robots distribués dans les divers réseaux. Ces approches ne peuvent cependant obtenir des résultats optimaux, voir même des résultats intéressants en raison de la faible fréquence des mesures et de leur disparité avec le contexte des services. La justesse de ces mesures dépend énormément du protocole employé pour la réalisation qui doit mettre en avant des moyennes statistiques suffisantes. Par ailleurs, des mesures effectuées dans le contexte particulier d'une entité ne sont pas toujours applicables pour une autre entité. le processus deviendrait rapidement lourd si l'on cherche à obtenir des informations pertinentes à travers ce genre de procédés.

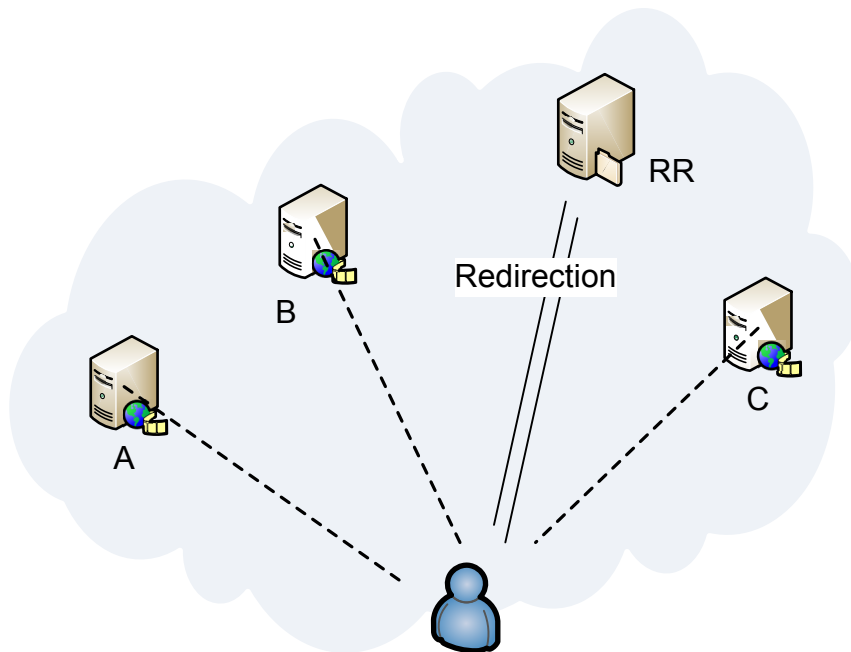


FIGURE 1.16 – Illustration de la diversité de chemins réseaux entre le client et les divers noeuds de distribution susceptible de le servir.

Le moyen le plus efficace pour atteindre une solution recevable à ce problème est de fournir aux opérateurs de CDN les informations nécessaires pour intégrer les conditions de transport du trafic dans l'équation de la sélection. Il s'agit à proprement parler de communiquer des informations relatives aux réseaux de transport aux opérateurs CDN et à bien d'autres services applicatifs susceptibles d'en tirer avantage en s'assurant de faire des choix alignés sur les contraintes en ressources des réseaux d'opérateurs qu'ils vont emprunter.

Cette approche semble sans doute évidente et cohérente pour atteindre les objectifs fixés à travers une telle collaboration. Il n'en reste pas moins nécessaire de poser la problématique à l'envers afin de chercher si les opérateurs réseaux ne pourrait pas tirer

d'avantage de profit d'une meilleure connaissance du trafic qu'ils devront acheminer en apportant des moyens plus adaptés ou bien des réponses plus satisfaisantes aux divers services applicatifs qu'ils s'apprêtent à collaborer avec. Ces mêmes services sont en effet en concurrence pour l'accès et l'utilisation des ressources réseaux si on suppose qu'ils appartiennent à la même classe de service et sont par conséquent traités de façon identique par les équipements des réseaux. En ayant à disposition des informations sur les contraintes en trafic auxquels ils sont ou seront exposés, les opérateurs réseaux seraient aptes à définir une stratégie plus efficace pour l'acheminement des données.

Dans un tel cadre, c'est une collaboration entre les acteurs réseaux et les acteurs des couches supérieures responsables de la gestion des services et applications proposées aux utilisateurs qui s'installe. Cette collaboration doit être régie par un système bien défini afin d'assurer une exploitation standardisée de ce procédé à une large échelle et de manière compatible avec les besoins et disponibilité des différents acteurs. La tâche de définir un tel système a été entreprise dans le cadre du projet européen ENVISION [UCoL09, MES⁺11a], intégrant les contributions des travaux de cette thèse. Les prochaines sections apporteront des éléments de réponses concernant l'architecture fonctionnelle du système proposé, l'interface de collaboration et son application dans le contexte du service de distribution de contenu CDN.

1.4.2.2 L'intérêt pour les applications P2P

Les applications P2P ont fait l'objet de travaux d'optimisation considérables qui traitent d'aspects variés comme les protocoles de signalisation et de transport, l'architecture des "overlay" P2P ou encore la sélection de pairs. Cependant, les applications P2P agnostiques des conditions des réseaux de transport induisent une charge de trafic toujours perfectible dans un contexte de transport de plus en plus tendu. Pour palier à ce problème, une nouvelle approche qui prône plus de communication entre les applications P2P et les réseaux qu'elles utilisent a émergé dans le but d'optimiser le transport des données. Les travaux réalisés par le projet *P4P* [XKSY, XYK⁺08] ou par le projet NAPA-WINE [ABB⁺10, BLM⁺11] et les scénarios P2P étudiés par le groupe de travail ALTO à l'IETF [SB09] intègrent cette voie d'optimisation. Toutefois, ils sont loin d'apporter les solutions les plus optimisées. Néanmoins, l'intérêt de cette approche est d'assurer une meilleure gestion du transport des données générées par les applications P2P qui continue de représenter une partie non négligeable du trafic Internet comme en témoigne divers études [Gil12, GPSH12]. Cette approche repose sur une caractéristique des systèmes P2P qui est la redondance des ressources. En effet, la capacité de différentes entités d'offrir les mêmes ressources ou services conduit au besoin d'effectuer des choix pour retenir parmi une liste de pairs, ceux qui vont fournir les ressources dont a besoin chaque client. Une interface de collaboration entre applications P2P et réseaux de transport pourrait améliorer ce processus de sélection de pairs en intégrant une dimension **transport** dans l'équation. En effet, la considération d'informations relatives aux performances du réseau et aux contraintes des applications permet de sélectionner les pairs qui offriront les meilleures performances pour un client P2P tout en ménageant

le réseau de transport. Ainsi, connecter des pairs localisés dans la même zone ou le même domaine par exemple permet non seulement de ménager des liens inter-domaine ou de transit mais aussi de raccourcir le délai de réponse éventuellement entre les pairs. Une interface de collaboration pourrait ainsi permettre d'améliorer les performances de transport pour les applications P2P tout en parvenant à une meilleure utilisation des ressources réseau. Dans cette optique, on peut espérer d'une collaboration plus poussée entre les applications et les réseaux, qui leur permet d'exposer leurs contraintes et déclarer leurs besoins, de fournir une solution optimale pour la gestion du trafic P2P.

1.4.2.3 Impacts stratégiques et financiers pour les opérateurs réseaux

Dans un contexte de plus en plus concurrentiel et une conjoncture économique tendue, toute technique d'optimisation n'a bien sûr d'intérêt pour les opérateurs de télécommunications que si elle dispose d'impacts non négligeable sur les aspects stratégiques et financiers.

Impacts stratégiques Dans le cadre de sa consultation publique sur la neutralité du net [RH12], la commission européenne a mandaté l'organisme de régulation européen "*Body of European Regulators of European Communications*" (BEREC) à lancer des investigations sur les pratiques de gestion du trafic par les opérateurs de télécommunications. Les résultats ont prouvé qu'environ 20% des fournisseurs d'accès fixe limitent le trafic de certains services comme le P2P. Les restrictions sont plus sévères pour les clients mobiles. Par ailleurs, les politiques de gestion de la congestion dans le réseau peuvent être identiques pour tout type de trafic ou différenciée pour favoriser un fournisseur de service particulier comme *Facebook*. Certaines de ces pratiques sont dans le viseur de la commission européenne. En conséquence, une gestion optimisée du trafic dans le réseau est nécessaire pour faire face à l'interdiction des restrictions. Une collaboration avec les applications P2P peut être déterminante dans la stratégie des opérateurs pour lever les restrictions qu'ils pratiquent tout en contrôlant l'augmentation de la charge sur leurs réseaux.

Impacts financiers La croissance du trafic internet en général et du trafic vidéo en particulier est un fait sur lequel tout le monde s'accorde aujourd'hui. Il n'y a qu'à voir les statistiques de *Youtube* dont je cite, quatre milliards de vidéos regardées chaque jour, plus de 60 heures de vidéos envoyées à Youtube par minute, plus de 800 millions d'utilisateurs distincts par mois, etc. En conséquence, cette croissance oblige les opérateurs à revoir continuellement le dimensionnement de leurs réseaux et à opter pour des technologies offrant de meilleurs débits pour faire face. Les investissements concédés peuvent atteindre plusieurs milliards d'euros par an et concernent toutes les infrastructures réseaux car la qualité de service et donc indirectement l'attractivité des opérateurs aussi qui en dépend. Nous pouvons espérer en conséquence qu'une meilleure gestion du trafic pourrait ralentir voir reporter certains investissements en contenant la charge de trafic à des niveaux gérables par les infrastructures existantes. Bien que les investissements dans les réseaux d'accès fixes ou mobiles ne soient pas concernés par notre proposition,

les arguments financiers sont toujours valables surtout que l'approche collaborative que nous soutenons ne requiert pas d'investissement de départ.

Par ailleurs, les opérateurs réseaux sont de plus en plus enclins aujourd'hui à signer des accords de *peering* avec des fournisseurs de services de distribution de contenus, des accords appelés aussi "*Content Peering*". Même si ces accords permettent aux opérateurs de réaliser des économies sur les coûts de transit, les bénéfices à long terme ne sont pas en leur faveur [Kro11]. En effet, d'un côté, une partie des revenus qui parvenait des coûts de transit facturés pour aux fournisseurs de contenus est perdue. D'un autre côté, des géants du web comme Google qui contrôle Youtube, dont on estime sa part dans le trafic mondial en transit entre 10 et 20%, et qui dispose d'infrastructure à l'échelle mondiale, se donne des atouts pour faire pression et négocier ses coûts d'acheminement. Au final, les opérateurs réseaux sont confrontés à une augmentation du trafic accompagnée d'une baisse des revenus. Dans un contexte concurrentiel, savoir optimiser la gestion de ses ressources tout en améliorant la qualité de service pourvue est un facteur déterminant. La technique que nous proposons s'inscrit dans cette optique pour permettre aux opérateurs de collaborer plus efficacement avec les fournisseurs de contenus afin d'assurer une meilleure utilisation des ressources du réseau. Les investissements peuvent être plus ciblés en conséquence et la rentabilité améliorée.

1.5 Conclusion

Nous avons présenté dans ce chapitre les travaux qui ont été entrepris dans les milieux académiques et industriels pour apporter une réponse à une sérieuse problématique de trafic engendrée par plusieurs facteurs dont l'émergence de poids lourds de l'Internet, la tendance multimédia des contenus, la démocratisation des terminaux connectés et l'expectative croissante des usagers. Nous avons analysé les propositions existantes et élaboré des synthèses autour des architectures définies, des interfaces utilisées et des techniques appliquées. Nous avons par la suite avancé les difficultés d'emploi de certains de ces systèmes et l'insuffisance d'autres. Puis nous avons présenté de réelles motivations qui nous ont amené à envisager une nouvelle approche plus efficace et plus flexible qui sera présentée dans le prochain chapitre.

Chapitre 2

L'approche collaborative entre réseaux et services

Nous avons présenté dans le précédent chapitre les différents mécanismes qui ont été proposés pour fournir des moyens capables d'interfacer les fonctionnalités des réseaux et des services. L'analyse des limites de certains de ces systèmes en termes de complexité, d'inadéquation et d'insuffisance, nous a amené à envisager une nouvelle approche plus efficace et plus flexible. Dans ce chapitre, nous introduisons l'architecture de notre système visant à permettre une collaboration efficace entre réseaux et services. Nous présenterons une interface simple et adaptée de conception permettant le support de divers services de collaboration. Enfin, nous détaillerons différentes techniques que nous proposons avec différents niveaux de collaboration et expliqueront leur mise en exploitation pour différents types d'applications.

2.1 L'interface de collaboration entre réseaux et applications

Nous avons abordé dans la section précédente les contours d'une nouvelle approche d'optimisation basée sur des interfaces de collaboration entre services et réseaux. Cette voie doit nous permettre d'apporter des solutions à la problématique d'augmentation de charge dans les réseaux et de besoins croissants en QoS tout en dépassant les limites et améliorant leur rendement des techniques existantes présentées dans la section 1.4.1.

Nous exposerons dans les prochaines sections de ce rapport notre contribution dans la recherche d'une solution à ces problématiques. Nous présenterons les techniques que nous proposons qui reposent sur une approche efficace et simple afin de limiter la complexité de mise en application et assurer un passage à l'échelle pour des réseaux et des services de natures variées. Nous commencerons dans cette section par énoncer les exigences pour la définition d'un tel système, puis nous introduirons l'architecture que nous avons conçu pour notre système et exposerons par la suite l'ensemble des exigences auxquelles

elle répond. Nous aborderons enfin les services fournis par l'interface de collaboration et leur mise en application.

2.1.1 Les exigences de l'approche collaborative

L'approche collaborative que nous avons évoquée plusieurs fois dans les sections précédentes vise à assurer la gestion des contraintes et des besoins des différents acteurs de transport et de service de façon conjointe et optimale. Nous énoncerons ci-dessous un ensemble de critères qui posent la problématique que nous nous sommes proposés de traiter.

Le support de multiples types services Le système proposé doit être indépendant d'un service particulier et assurer un support générique pour divers types d'applications comme les CDNs, P2P ou Cloud par exemple. En effet, la pertinence des résultats obtenus dépendra de la capacité du système à optimiser le transport de la plus grande partie du trafic possible dans le réseau quelque soit sa source.

La gestion commune du trafic La gestion du trafic de plusieurs services applicatif est une procédure qui requiert un traitement bien particulier. La problématique se dégage de la façon dont est gérée la collaboration avec chaque service. Des applications de natures différentes et avec des besoins distincts ne peuvent bénéficier du même traitement. Par contre, le trafic est généralement traité de façon identique dans le réseau (on ne s'intéresse pas aux mécanismes de différenciation de qualité de service qui sont rarement utilisés pour des services "*over the top*"). Ces constatations nous amène à définir un système qui fournit un traitement particulier et optimisé pour chaque service à partir d'une gestion commune de l'ensemble des contraintes et besoins de tous les acteurs.

Une interface simple et flexible L'approche collaborative que nous proposons s'adresse tout aussi bien aux services proposés par les opérateurs (par exemple un système de distribution de contenus CDN propre) qu'aux services *overlay* développés généralement sur la base de technologies web dissociées des normes et protocoles réseaux et Télécom. L'utilisation de l'interface doit rester accessible en garantissant une simplicité d'intégration et d'exploitation et une flexibilité vis-à-vis des différences de contexte. Toute la complexité doit à ce titre épargnée aux services applicatifs.

Une optimisation fiable et efficace L'intérêt de la collaboration proposée réside principalement dans l'importance des gains réalisés. L'intelligence du système doit disposer de procédures d'optimisations qui soient à la fois fiables pour garantir l'intégrité des réseaux de transport et du trafic mais aussi efficaces pour fournir des résultats probants en termes d'utilisation des ressources réseaux et de l'amélioration de la QoE des services applicatifs.

2.1.2 L'architecture du système de collaboration

L'adoption de l'architecture que nous proposons est tributaire de son indépendance d'une technologies ou infrastructure particulière. Cette position nous offre une grande flexibilité dans l'utilisation au détriment d'une intégration plus exigeante.

2.1.2.1 Présentation de l'architecture

Dans cette section, nous présentons l'architecture fonctionnelle globale de notre système. Nous ne précisons pas si les fonctions définies sont centralisées au sein d'une entité ou distribuées (au sein de certaines entités comme les utilisateurs finaux). Les recommandations d'intégration dans des architectures organiques dépendront de fait de la nature du système ou du service. L'architecture a été spécifiée pour répondre aux exigences de la collaboration prévue entre les entités participant à la distribution des contenus. Un ensemble de blocs fonctionnels capitaux pour réaliser cette collaboration a été identifié et les interfaces entre eux correctement définies. En ce qui concerne le réseau de transport sous-jacent, étant composé de Systèmes Autonomes (AS) opérant de manière autonome leurs ressources, nous avons défini un niveau d'abstraction articulé autour d'une fonction de gestion des ressources sous la responsabilité de chaque opérateur réseau souhaitant coopérer avec les applications utilisant ses ressources. La collaboration a été identifiée comme une exposition des informations pertinentes quant à l'utilisation des ressources et réciproquement la mise à disposition de services. Un tel niveau de collaboration exige l'implémentation de mécanismes de sécurité pour offrir des facilités d'authentification, d'autorisation et de comptage si nécessaire. Quant aux applications et services impliqués dans cette collaboration, des fonctions de partage d'informations sont nécessaires. D'autres services capables d'offrir des fonctionnalités d'amélioration de la qualité d'expérience ressentie par les usagers, comme par exemple l'adaptation du contenu, peuvent être fournies par le service applicatif, le réseau ou un fournisseur de services tiers.

2.1.2.2 Présentation des modules fonctionnels

Il est intéressant de voir qu'une séparation des modules fonctionnels s'impose naturellement avec des fonctionnalités implémentées au niveau réseau et d'autres au niveau du service applicatif. Nous ne nous attarderons que sur les fonctions pertinentes à notre modèle. En effet, certains volets comme la sécurisation de l'interface font intervenir des techniques bien rodées et largement déployées. De ce fait, les blocs fonctionnels "Authentication, Authorisation and Accounting (AAA) ne seront pas présentés. Par ailleurs, le module fonctionnel "Service Control", gestionnaire des services tiers disponibles au soutien de l'application comme des services de stockage ou de cache, ne sera pas traité non plus. Nous aborderons d'abord la présentation des modules fonctionnels du service applicatif pour ensuite s'atteler aux fonctions définies pour le réseau et par ordre d'importance croissant pour aboutir enfin au module implémentant l'intelligence du système.

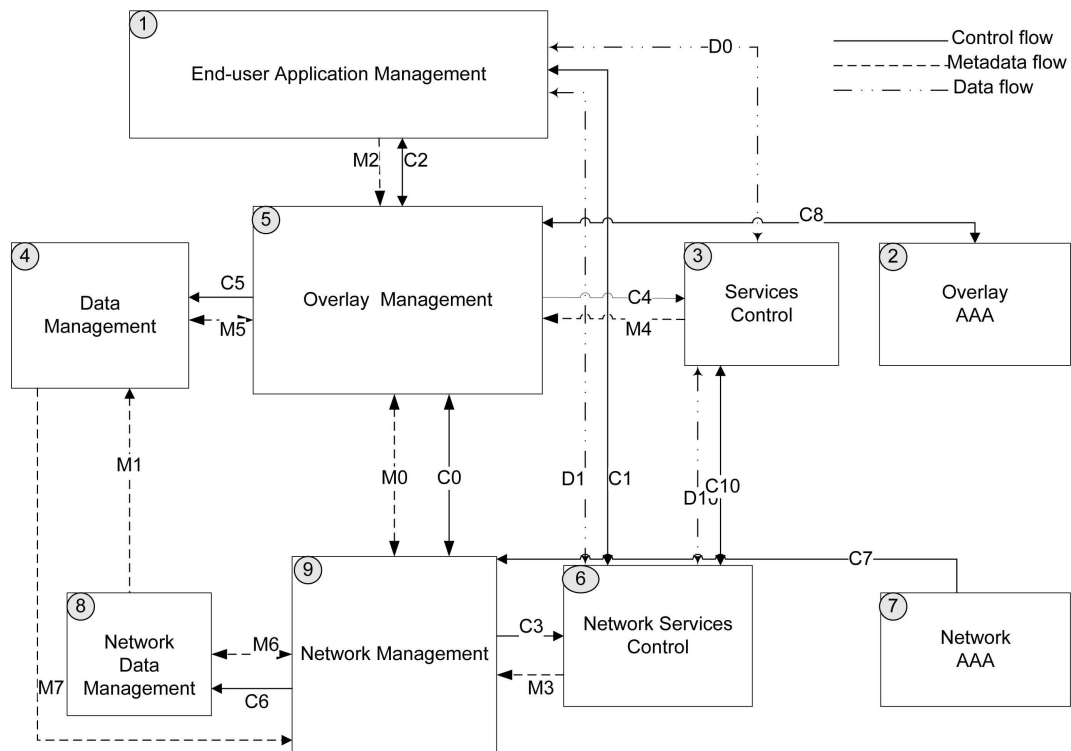


FIGURE 2.1 – L'architecture représente les divers blocs fonctionnels nécessaires au système de collaboration entre réseaux et applications avec les divers interfaces permettant l'acheminement des requêtes et données entre eux.

EUAM Le "End-User Application Management", gestionnaire de l'application utilisateur est le module fonctionnel responsable de la gestion des fonctionnalités de l'application exécutées côté utilisateur final. Ce module diffère d'un type d'application à un autre. En l'occurrence, les applications distribuées telles les applications pair-à-pair "P2P" sont exposées à des contraintes différentes des modèles client-serveur. Ce module fonctionnel met en oeuvre les procédures permettant à l'utilisateur de se connecter au service applicatif "overlay", de transmettre et réceptionner des informations et de négocier avec le gestionnaire du service applicatif le contenu d'intérêt, les préférences de l'utilisateur, les capacités du terminal, la contribution en ressources pour le service applicatif (stockage de données, adaptation de contenus, redistribution...). Ce module est responsable aussi de la bonne gestion de la connectivité réseau notamment en surveillant la qualité de la session applicative (retard, taux de perte...) et en contrôlant la connectivité du terminal.

ODM Le gestionnaire des données du service applicatif "Overlay Data Management" est le module fonctionnel responsable de la gestion des données collectées par l'application et qui sont relatives aux réseaux de transport, au service applicatif, aux utilisateurs, etc. Ces informations s'intègrent dans la sphère du contexte global de l'application et sont nécessaires pour l'optimisation du fonctionnement du service applicatif, plus particulièrement concernant la gestion du trafic applicatif généré dans les réseaux de transport. Ces informations incluent :

- Informations sur les réseaux relatives à l'état du réseau, conditions de transport, disponibilité des ressources et des services (comme le multicast), localisation des infrastructures mises à disposition (noeud source pour le multicast ou serveur de cache par exemple), etc.
- Informations relatives à l'application et qui incluent les capacités des noeuds applicatifs, la disponibilité de leurs ressources, la disponibilité de services tiers, les conditions d'utilisation, etc. Ces données intègrent aussi des renseignements sur les contenus disséminés, leur disponibilité dans les divers entités de service, les paramètres de performance de l'application, etc.
- Informations sur les contenus incluant les méta-données concernant tout élément de contenu distribué par l'application, comme le format source d'origine, le producteur du contenu, la nature, les restrictions de distribution, etc.

Ce module assure le stockage persistant des informations exposées ci-dessus et collectées depuis différents autres modules fonctionnels.

OM Le gestionnaire du service applicatif, "Overlay Management", est le module fonctionnel responsable des décisions relatives au fonctionnement de l'application et aux interactions avec les acteurs impliqués dans l'optimisation du service qui sont les utilisateurs et les réseaux de transport. Cette fonction décide de l'utilisation des ressources et services disponibles pour l'application afin d'assurer un fonctionnement optimal aligné sur les contraintes exposées par les réseaux de transport. Il s'agit par exemple de décider de l'allocation de certains noeuds de distribution à tel ou tel contenu en fonction

de sa popularité, de demander l'activation de certains services mis à disposition par les opérateurs réseaux ou des fournisseurs tiers ou d'inciter les utilisateurs à contribuer à la distribution des contenus à travers des connexions pair à pair, etc. L'objectif étant de trouver un équilibre satisfaisant entre la qualité d'expérience ressentie par les utilisateurs et le coût associé à l'utilisation des ressources et services disponibles tout en veillant que ces actifs sont exploités de façon optimale.

Les modules fonctionnels qui seront présentés par la suite sont implémentés par les opérateurs réseaux. Certains modules présentent des similitudes avec les modules fonctionnels du service applicatif, à travers des fonctionnalités similaires.

NDM Le gestionnaire des données du réseau, "Network Data Management", est un module responsable de la collecte et la gestion des données exploitées pour l'optimisation de l'utilisation des ressources et services réseaux. Il fournit des fonctions pour la collecte, le traitement et le stockage des données, la gestion de l'accès aux informations par les autres modules fonctionnels, la gestion de la mise à jour des informations, des politiques de surveillance et notifications, etc.

NSC Le contrôleur des services réseau, "Network Service Control" est l'entité fonctionnelle responsable de la gestion des procédures d'exécution et surveillance des services réseau sur la base des règles définies par le gestionnaire du réseau. Le contrôleur implémente un certain nombre de fonctionnalité incluant la réservation des ressources réseaux pour le trafic du service applicatif, la gestion de la transmission des données en multicast, la surveillance des conditions de transport dans le réseau, la gestion des serveurs de cache, de stockage ou d'adaptation de contenu disponible dans le réseau, la surveillance du fonctionnement des services réseau, la remontée d'information au gestionnaire réseau, etc. Le contrôleur est implémenté par les opérateurs réseaux pour le support des divers fonctionnalités offertes par le réseau.

NM Le gestionnaire réseau, "Network Management", est le module fonctionnel implémenté par les opérateurs réseaux chargé du traitement des demandes du service applicatif dans le cadre de cette coopération prévue par l'interface CINA. Il s'agit par ailleurs d'une entité responsable de la coordination de l'ensemble des actions à mettre en oeuvre dans le réseau. Ce module implémente l'intelligence du système au niveau du réseau pour le traitement, la décision et le contrôle des actions entreprises à ce niveau. De ce fait, il assure plusieurs fonctions :

- La collecte des données depuis divers modules fonctionnels du réseau ou du service applicatif,
- L'analyse et le traitement des informations,
- L'exécution des algorithmes de décision,
- Le contrôle des modules réseau,
- La gestion de l'interface avec les services applicatifs.

Les algorithmes de décision déployés au sein de cette entité fonctionnelle représentent les procédures intelligentes qui définissent les actions à entreprendre pour optimiser le fonctionnement du système et plus particulièrement du réseau. Les décisions sont définies en fonction de l'ensemble des informations relatives aux ressources, services et politiques réseau et celles similaires relatives aux services applicatifs.

2.1.3 L'interface CINA : Collaboration Interface Between Networks and Application

2.1.3.1 Présentation

L'interface de communication entre les réseaux et les applications, CINA, est une composante centrale du modèle de collaboration que nous avons développé. Divers travaux ont été menés conjointement pour la spécification de cette interface afin de répondre aux divers besoins fonctionnels, techniques, et de sécurité. Une partie des travaux était dédiée à la spécification du protocole de communication qui sera utilisé par les acteurs afin de préciser les échanges permettant les négociations et les communications d'informations de part et d'autres. Une autre partie des travaux a été dédiée à la spécification des mécanismes nécessaires à la découverte de la disponibilité d'outil de collaboration dans chaque réseau. Il est nécessaire de retenir qu'il en va du souhait de chaque opérateur réseau comme de chaque gestionnaire d'application quant à l'adhésion à ce modèle collaboratif. En conséquence, des mécanismes permettant aux applications de découvrir si l'opérateur réseau adhère au système de collaboration sont nécessaires. Par ailleurs, d'autres travaux ont été consacrés à la spécification de mécanismes pour la sécurisation de l'utilisation de l'interface intégrant des procédures d'authentification mutuelle, de contrôle d'accès aux informations, de chiffrement des échanges, etc. Ces spécifications auxquelles nous avons pris part sont présentées en détail dans les différents livrables du projet ENVISION [BM11a, BM11b]. Nous nous intéresserons spécialement dans ce document au protocole de communication entre les modules réseaux et les modules applicatifs. Les différents volets des travaux comme la découverte de l'interface et sa sécurisation sont traités dans les livrables *D3.1* et *D3.2* [BM11a, BM11b].

2.2 L'approche de collaboration passive des opérateurs réseaux

2.2.1 Introduction de l'approche

Cette approche est fondée sur une collaboration passive des opérateurs réseaux qui sont sollicités par les services applicatifs afin de fournir des informations sur les conditions de transport dans leurs réseaux. Les opérateurs réseaux se trouvent amenés à collaborer avec les applications "overlay" sans pour autant intervenir dans les prises de décisions relatives à l'aiguillage du trafic applicatif. Par ailleurs, cette passivité se traduit aussi par l'ignorance par les opérateurs réseaux des contraintes en trafic que vont exercer les applications sur leurs réseaux. Il s'agit néanmoins d'une approche bien éta-

blie qui pourrait correspondre à des contextes spécifiques où d'autres types d'approches ne pourraient être exploitées en raison de contraintes stratégiques, économiques ou techniques.

Techniquement, les services applicatifs sont amenés à faire des choix qui ont des conséquences différentes sur les réseaux de transport. Ces choix peuvent avoir une dimension spatiale, par exemple rediriger un client vers un serveur particulier pour les CDNs ou bien établir une liste de pairs pour un nouvel arrivant dans les applications "P2P". Chaque choix se traduira par un chemin réseau partiellement différent ce qui implique l'utilisation de ressources réseau sous conditions diverses. Les choix peuvent aussi avoir une dimension temporelle quand il s'agit de définir une planification dans le temps de l'utilisation des ressources réseaux. Un cas d'étude concret pourrait cibler les transferts de données entre Data-Centers (DCs) qui peuvent être réalisés pendant des tranches horaires bien spécifiques où les réseaux offrent plus de disponibilité, moyennant parfois des coûts moindres. Pour soutenir les applications dans leur prise de décision, deux besoins fonctionnels se dégagent :

1. La localisation des entités applicatives dans le réseau : Il s'agit de permettre aux applications de déterminer la localisation de leurs entités clients, serveurs, etc.
2. Les conditions de transport entre les entités afin de faire un choix optimal au regard des ressources réseau disponibles

2.2.2 Application de l'approche

Dans cette section, nous aborderons l'application de cette approche de collaboration passive au cas particulier des services CDNs. Ce choix est légitime au vu de la proportion de trafic CDN qui transite aujourd'hui dans les réseaux et la place qu'occupe les grands opérateurs CDNs dans le classement des principaux acteurs du trafic internet.

2.2.2.1 Le contexte réseau pour les services CDNs

Le contexte réseau a une incidence directe sur l'élaboration du schéma technique de collaboration entre les réseaux et CDNs. En effet, l'adhésion à ce modèle de collaboration requiert l'étude des architectures et topologies réseaux et leurs implications sur le transport du trafic. A titre explicatif, nous ne citerons pour les clients du réseau fixe, abonnés à l'ADSL par exemple, que les équipements les plus pertinents. Ils sont rattachés à des DSLAMs, connectés à des noeuds de périphérie (NE), connectés à des noeuds multimédias (NM), connectés à des noeuds de concentration (NC), connectés à des noeuds régionaux (NR), connectés à des noeuds de transit (NT). En supposant que les noeuds de distribution d'un CDN sont hébergés à un niveau plus haut que les équipements NRs dans le réseau, on constate que le trafic CDN des abonnés ADSL qui sont rattachés au réseau coeur à travers un équipement NR, emprunte un chemin réseau en deux tronçons. Un premier tronçon entre les clients et l'équipement NR qui est statique quel que soit le noeud de distribution vers lequel ils seront redirigés. Un

deuxième tronçon entre le NR et les noeuds de distribution qui présentent une diversité de chemin. On peut tirer deux enseignements de cette étude :

1. La différence en qualité de service relative au chemin réseau entre les clients CDNs rattachés au même équipement NR intervient en aval de cet équipement en présence d'une diversité de chemins d'où la possibilité de contenir le problème d'optimisation du transport des données à ce tronçon dynamique.
2. Les entités rattachées au même NR peuvent être considérées comme une seule entité applicative ou bien regroupées ensemble vu que le tronçon statique en amont du NR n'a pas d'incidence sur l'optimisation du transport.

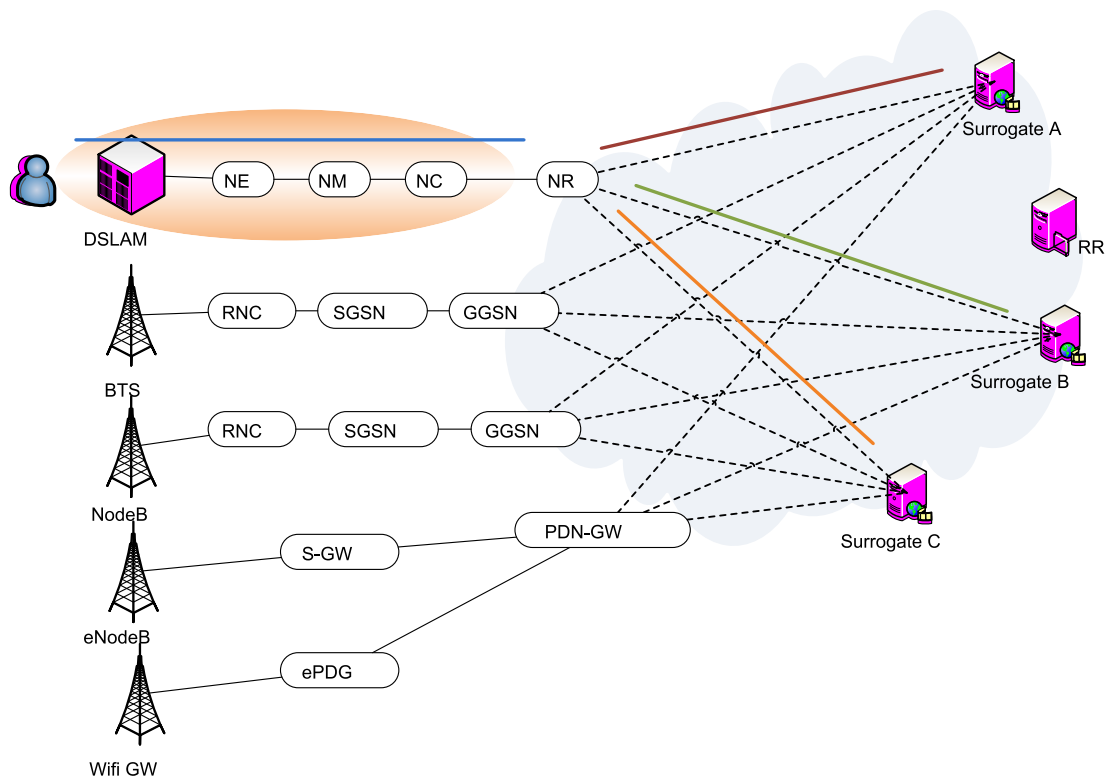


FIGURE 2.2 – La figure illustre l'existence d'un tronçon statique sans incidence par rapport à l'optimisation du transport pour différents types de réseaux d'accès. En conséquence, le problème peut être contenu dans la partie du réseau présentant une diversité de chemin.

2.2.3 ALTO pour interface

L'interface de communication nécessaire à la mise en pratique de cette approche de collaboration passive doit assurer les besoins fonctionnels définis dans les sections précédentes. Le protocole ALTO [APY12] défini par le groupe de travail du même nom à l'IETF [ALT12] propose aujourd'hui une interface compatible avec ces besoins.

Cette même interface a constitué une rampe de départ au projet ENVISION [UCoL09] pour la spécification de l'interface de collaboration CINA, qui proposent des services supplémentaires permettant des fonctionnalités avancées dont une partie a été proposée en standardisation au groupe ALTO [NJWB⁺12]. Nous avons introduit le protocole ALTO dans la section 1.1.2 qui commence à la page 27.

2.2.4 Les différentes classes d'information pour l'interface ALTO

L'interface ALTO présentée dans la précédente section fournit les services escomptés pour soutenir les opérateurs CDN soucieux de prendre en considération le contexte du réseau de transport dans le problème de sélection du noeud de distribution pour leurs clients. ALTO, étant avant tout un groupe de travail en standardisation, limite ses efforts à la spécification des protocoles, moyens et services qui permettront la communication des informations des opérateurs réseaux aux fournisseurs de services notamment les opérateurs CDNs. ALTO ne fournit ni réponse ni recommandation quant aux questions relatives aux approches adoptées pour définir les cartes du réseau et les coûts liés, ni ne traite la pertinence des informations exposées. Nous rappelons que notre intérêt est porté sur l'optimisation de l'utilisation des ressources réseaux et l'amélioration de la qualité des services applicatifs en exploitant les informations de contexte relatives aux réseaux de transport. Dans ce cadre, nous n'aborderons pas d'approches qui peuvent se baser ou tenir compte de contraintes ou accords financiers ou toute information de nature autre que celles liées au réseau. Nous analyserons dans les sections suivantes différentes classes d'informations que les opérateurs réseaux pourraient exposer aux CDNs à travers l'interface ALTO. Ensuite, nous proposerons une nouvelle technique de calcul des coûts basée sur l'état d'utilisation des liens du réseau. Cependant, nous commencerons par aborder les défis rencontrés par les opérateurs réseaux pour fournir les services ALTO.

2.2.4.1 Les défis inhérents à la collecte et exposition des informations

Les opérateurs de réseaux qui implémentent l'interface ALTO fournissent un ensemble de services autour des informations exposées par les cartes du réseau et les matrices de coûts. Ces informations sont au préalable collectées et traitées par l'opérateur réseau par ses propres moyens.

La collecte des informations pour la construction des cartes du réseau présente moins de contraintes que la collecte des informations pour le calcul des matrices de coûts. En effet, les entrées de la carte (par exemple sous-réseaux ou PoPs) sont moins fréquemment mis à jour. Deux techniques principales pourraient être utilisées pour collecter ces informations. La première s'appuierait sur les informations fournies par les protocoles Exterior Gateway Protocol (EGP) pour un contexte inter-domaines ou Interior Gateway Protocol (IGP) intra-domaine comme Open Shortest Path First (OSPF) ou Intermediate System to Intermediate System (IS-IS) qui construisent des représentations logiques de la topologie du réseau. La seconde reposerait sur le plan de commande

pour récupérer les informations directement depuis les équipements du réseau. Intervient ensuite la construction des cartes du réseau. Cette étape est régie par le choix du niveau de granularité des informations qui influent directement sur les performances du système. En effet, localiser par exemple plusieurs sites de noeuds de distribution dans un même PID revient à ignorer tous les liens réseaux sur les chemins de ces serveurs appartenant à ce PID au moment de la sélection. En contre-partie, une granularité très fine incomberait un volume de données considérable sans toutefois apporter plus d'efficacité. Un exemple précis est la présence d'une portion invariable pour tous les chemins entre le client et les noeuds de distribution 1.16. Les liens réseaux constituant cette portion peuvent être ignorés de la carte réseau sans affecter la performance du système. Il s'en suit que la construction des cartes du réseau doivent obéir à certaines règles et contraintes notamment celles fixées par la politique de l'opérateur en termes de divulgation d'informations. La construction de cartes réseau spécifique à chaque service applicatif constitue sans doute la meilleure façon de faire. Elle nécessite la connaissance de l'emplacement des entités applicatives, et plus particulièrement des entités serveur comme les noeuds de distribution pour les CDNs. Néanmoins, pour s'adresser à un grand nombre de services applicatifs avec des architectures différentes comme le P2P, l'opérateur réseau devrait entamer une étude réfléchie de la topologie de son réseau afin de définir le niveau d'intervention sur le trafic réseau dont découlerait le niveau de spécification des PIDs pour une ou plusieurs cartes génériques. Les opérateurs réseaux doivent aussi être prudents vis-à-vis des attentes des fournisseurs de services en exposant des informations d'intérêt commun bénéfiques aux deux acteurs.

La tâche la plus ardue qui guette les opérateurs réseaux reste la construction des matrices de coûts. La complexité dépend bien sûr de la granularité de la carte du réseau et donc le volume de données à traiter mais surtout de la nature des informations à collecter pour calculer les coûts associés. En effet, quand certaines métriques comme le nombre de sauts sont quasi-statiques, d'autres comme la latence du chemin réseau de bout en bout doivent être mises à jour fréquemment pour garantir une certaine précision. Et encore la latence n'est pas toujours la métrique la plus appropriée dans des contextes de gros volumes de trafic. A l'échelle d'un réseau d'opérateur national ou international, cette collecte d'informations requiert une durée de temps variable de quelques minutes à plusieurs dizaines suivant la nature des informations et l'étendue du réseau. Par ailleurs, les opérateurs réseaux peuvent aussi limiter la fréquence de mise à jour des informations pour se conformer à une politique de confidentialité par exemple. En parallèle, le protocole ALTO de conception ne prône pas un transfert de connaissance en temps réel, quoique cette contrainte soit mineure. Cependant, la pertinence des informations exposées vis-à-vis du profil de service applicatif représente un facteur majeur dans la performance du système. En effet, les contraintes diffèrent d'un service à un autre. Les services de diffusion vidéo par exemple sont plus sensibles à la disponibilité d'une bande passante suffisante et constante. En revanche, des services de profil différent, communément appelés accélération du web, offrent la possibilité de télécharger des petits fichiers, généralement des pages, des images ou des petits volumes d'un site internet par exemple depuis un autre noeud que le serveur d'origine. Ces

transferts de données se font en une fraction de temps négligeable par rapport à la durée d'une session de diffusion par contre le temps d'accès aux données est le facteur principal de performance. L'étude des besoins et contraintes de chaque profil de service peuvent amener à l'utilisation de métriques différentes en adéquation avec le profil de chacun, par exemple la disponibilité de bande passante pour la diffusion vidéo et le délai ou la proximité pour l'accélération web. Ces aspects résument les principaux défis inhérents à la collecte et exposition des informations et qui doivent être pris en considération pour l'analyse des différentes catégories d'informations.

2.2.4.2 La classe d'informations distance

La matrice de coûts que peut fournir un opérateur réseau à un fournisseur de service peut être calculée à partir des informations de distance entre les PIDs définis par la carte du réseau. Dans ce cas, le coût d'utilisation d'un chemin réseau entre deux PIDs représente la distance en nombre de sauts ou équipements réseaux entre eux. La figure 2.3 illustre un exemple d'utilisation de cette métrique pour un CDN disposant de trois noeuds de distribution dans les PIDs 1, 2 et 4. La fonction Request Routing (RR) qui contrôle la redirection du client dans le PID 6 vers un serveur approprié est localisée à travers la carte du réseau dans le PID 3. Dans ce cas de figure, les distances fournies par le réseau entre le client et les divers serveurs sont exposées dans le tableau 2.1. Ces distances permettent à l'opérateur CDN de sélectionner le serveur le plus proche du client et par conséquent espérer un délai de transport moindre dans le réseau. Dans l'exemple de la figure 2.3, deux serveurs sont à égale distance réseau du client. Dans ce cas ci, à charge de la fonction RR du CDN de se fier à d'autres critères de sélection pour les départager. Dans la réalité, cette métrique présente un certain nombre d'inconvénients car elle ne fournit aucune indication sur l'état du réseau de transport. En effet, les choix basés sur cette classe d'informations s'avèrent sous-optimaux voir contre productifs dans certaines situations concrètes.

Premier cas de figure Supposons dans cet exemple la présence d'un trafic élevé sur le lien entre les deux routeurs RN2 et RN3. A distance égale, les deux serveurs dans le PID 1 et PID 2 peuvent être choisis indifféremment (ou se voir partager la tâche). Cependant, il serait préférable de rediriger les clients dans le PID 6 vers le serveur A dans le PID 1 afin d'emprunter un chemin réseau moins utilisé du coup offrant une meilleure disponibilité aux trafic des clients CDN.

Deuxième cas de figure Dans ce cas de figure, supposons une charge élevée sur le lien entre les deux routeurs RN1 et RN2. Par conséquent, le trafic sur les chemins réseau vers les serveurs A et B en pâtiraient empruntant ce lien particulier. Comme la matrice de coût basée sur la classe information distance ne permet pas de fournir ce genre de détail, le contrôleur du CDN choisira un des deux serveurs pour répondre aux requêtes du client. Pire encore, en continuant à rediriger les clients vers ces serveurs, il peut contribuer à aggraver la situation. Si le CDN dispose d'outils capables de surveiller la qualité de son service, il pourrait détecter les désagréments que les clients seraient

susceptibles de rencontrer et par conséquent rectifier ses choix. Autrement, un certain nombre de procédures peuvent être lancées dans le réseau afin de réduire la charge sur les liens concernés. Il peut s'agir de l'activation du mécanisme "Explicit Congestion Notification (ECN)" qui affecte les débits de transfert ou bien du faire du "Traffic Engineering (TE)" pour l'utilisation d'une autre route pour une partie du trafic et dans les situations extrêmes, une suppression de paquet intervient au niveau des routeurs pour contenir la charge.

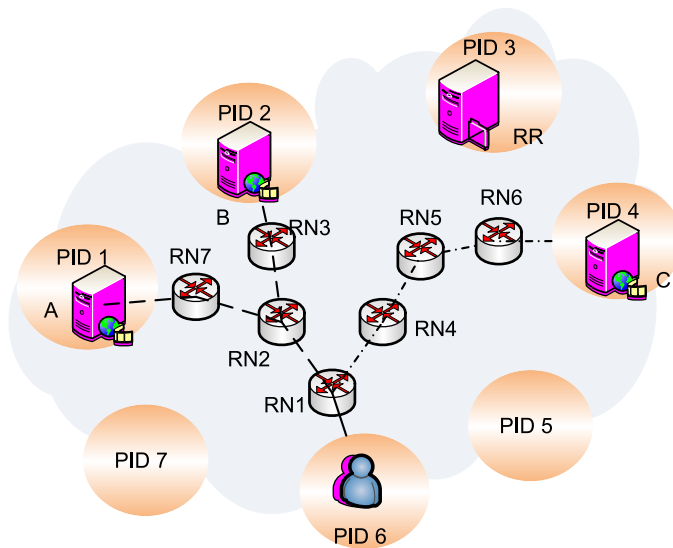


FIGURE 2.3 – Illustration de la métrique distance dans une carte du réseau

TABLE 2.1 – Distance entre le client et les noeuds de distribution

PID Client	PID	
	Ser- veur	Distance
6	1	3
6	2	3
6	4	4

Les scénarios précédents ont souligné l'insuffisance des informations de distance, agnostique de l'état des liens composant les chemins de bout-en-bout entre les PIDs, ce qui amènerait à des choix pas assurément performants pour les acteurs concernés. Par ailleurs, ces données doivent être prudemment conditionnées afin de ne pas divulguer des informations confidentielles sur la topologie réseau.

2.2.4.3 La classe d'informations latence

Le calcul de la matrice de coûts peut se baser sur les informations de latence de bout en bout entre les PIDs de la carte du réseau. Ces valeurs peuvent être récupérées par mesure du "Round-Trip Time (RTT)" entre les noeuds de bordure des PIDs directement ou par accumulation des délais de transmission et propagation entre noeuds adjacents du réseau. Cette classe est défavorisée par un certain nombre de handicaps relatifs à la nature des informations. Nous dégageons d'un côté un besoin d'assurer une mise à jour fréquente des données afin de fournir des mesures de précision faute de quoi la performance du système est négativement affectée. D'un autre côté, la latence réseau n'est pas un candidat approprié pour un nombre de profils de services, et plus particulièrement les services de diffusion vidéos, assujettis à de gros volumes de transfert. Nous considérons une latence réseau plus importante (généralement on ne dépasse pas une différence de quelques centaines de millisecondes dans les réseaux terrestres) comme un facteur secondaire comparé à la capacité d'acheminer le trafic. Concrètement, ce procédé de récupération d'informations de latence n'offre que des estimations approximatives sur l'état et la capacité des chemins réseau, dont l'efficacité est affectée par fréquence des mesures dans le temps.

2.2.4.4 La classe d'informations bande passante

La bande passante disponible entre les PIDs peut être considéré comme une classe d'informations susceptibles d'être fournie par les opérateurs réseaux. En surveillant les interfaces de leurs équipements réseau, des protocoles comme SNMP permettent aux opérateurs de récupérer la bande passante disponible entre les équipements adjacents. Cette classe peut fournir des indicateurs pertinents pour certains services sensibles à la disponibilité de bande passante si la fréquence d'actualisation des informations est acceptable. Elle permettrait une meilleure utilisation des ressources du réseau à travers une meilleure distribution de la charge du trafic entre les chemins réseaux disponibles pour l'ensemble des clients.

Malgré les avantages préfigurés par ce type de données réseaux, deux défis majeurs se dressent sur la voie de son déploiement. Le premier défi, de nature technique, consiste à calculer la bande passante de bout en bout sur le chemins réseau entre client et serveur. Les valeurs du maximum, minimum ou la moyenne des bandes passantes des liens qui constituent le chemin ne fournissent pas d'indications exploitables. Le maximum et la valeur moyenne peuvent conduire à une surcharge des liens avec des valeurs inférieures alors que la valeur minimale pourrait être celle d'un lien présent sur divers chemin et ne permettrait donc pas de départager entre eux. La figure 2.4 présente un cas concret qui illustre la difficulté d'usage des informations de bande passante. Le lien entre les équipements réseaux RN1 et RN2 dispose d'une bande passante disponible de 10 Gb/s , utilisée par les clients des PID 6 et 7. Entre RN2 et le PID 2, la bande passante est de 45 Gb/s et 20 Gb/s entre RN2 et le PID 4. Par contre, la matrice de coûts fournie au CDN ne délivre que les valeurs de bout en bout entre les PIDs (voir le tableau 2.2). Ces

informations ne permettent pas de réaliser des choix assurément efficaces dans toutes les situations. Un opérateur réseau ne peut indiquer aux services applicatifs qu'ils disposent de 10 Gb/s entre les PIDs dont les chemins empruntent ce lien car il revendiquerait 10 Gb/s pour chaque chemin. En même temps, il ne pourrait partager cette valeur entre les différents chemin au risque d'allouer plus de ressources au PID qui en a le moins besoin. Dans un tel contexte, il semble impossible de fournir ce genre d'information en l'état à des services applicatifs.

Enfin, il semble aussi très difficile qu'un opérateur réseau puisse exposer des données de cette nature au regard du degré d'importance et de confidentialité de ces informations. En dehors de ces propres services comme un CDN interne, nous doutons sérieusement qu'un opérateur réseau ait la volonté d'exposer ces informations.

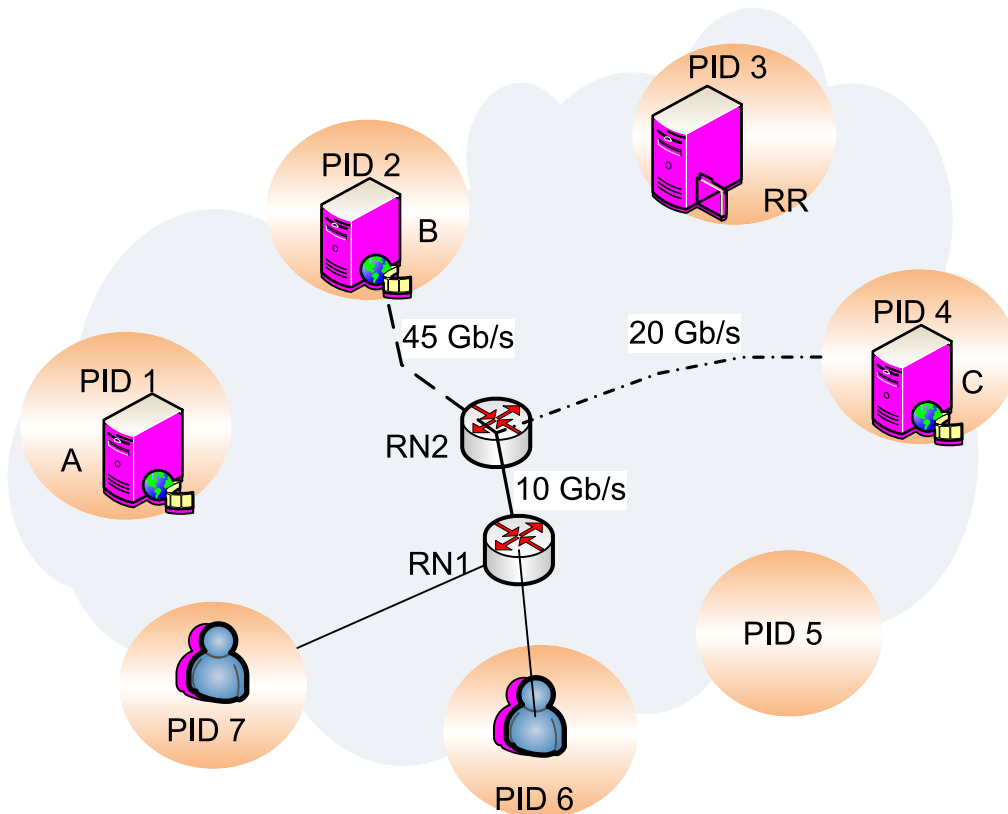


FIGURE 2.4 – Illustration de la métrique bande passante dans une carte du réseau

2.2.4.5 Autres classes d'informations

Les opérateurs réseaux sont susceptibles de se tourner vers d'autres classes d'informations pas forcément indicatrice des conditions de transport dans le réseau pour

TABLE 2.2 – Utilisation de la métrique bande passante pour une matrice de coûts

PID Client	PID Ser- veur	Min (Gb/s)	Max(Gb/s)	Moyenne (Gb/s)
6	2	10	45	27.5
6	4	10	20	15

construire des matrices de coûts à leur avantage. Nous retenons particulièrement les préférences unilatérales des opérateurs en dépit des disponibilités de ressources qui feraient que le trafic des services applicatifs serait acheminé via des chemins moins performants mais plus rentables pour les opérateurs réseaux. Le trafic inter-domaine via des noeuds de raccordements ou "peering" peut être assujéti à une politique plus favorable à l'opérateur dont découlerait une matrice de coûts soigneusement calculée. Il s'agit ici d'un axe de recherche qui favorise des critères différents des nôtres comme des objectifs financiers par exemple. Certains travaux en ce sens ont été menés dans le cadre du projet ENVISION. Nous invitons les lecteurs à lire les travaux exposés dans l'article suivant qui traite de l'optimisation des coûts financiers du trafic inter-domaine suivant une distribution spatiale et temporelle exploitant l'interface CINA [LMC⁺12]. Cette approche peut cependant conduire à des résultats contre-productifs si les objectifs des opérateurs réseaux et fournisseurs de services ne sont pas alignés en vue d'un compromis bénéfiques pour les deux acteurs.

2.2.5 Proposition d'une nouvelle classe d'informations

Bien que la liste des classes d'informations que nous venons d'évoquer n'est pas exhaustive, elle comprend néanmoins les principales métriques susceptibles d'être utilisées par les opérateurs réseaux. Par ailleurs, à notre connaissance, il n'y a pas d'autres types d'informations évoqués par la littérature ou par le groupe de travail ALTO à l'IETF pour être utilisés pour le calcul des matrices de coûts.

L'analyse des différentes classes d'informations que nous avons menée dans les précédentes sections a montré que bien que certaines métriques sont inutilisables en l'état, d'autres peuvent contribuer à l'amélioration de la performance globale des acteurs contributeurs de la collaboration sans toutefois garantir les meilleurs résultats. Pour remédier aux lacunes de ces classes d'informations, nous proposons une nouvelle approche pour calculer les valeurs des coûts exposés par la matrice de coûts à travers l'interface ALTO.

2.2.5.1 La classe d'information coût de transport

Nous définissons le coût de transport entre un PID source et un PID destination comme le coût d'acheminement des données sur le chemin réseau de bout en bout entre eux. Quand plusieurs chemins sont comparés, celui qui présente le coût le plus faible est

retenu pour acheminer le trafic et le client est en conséquence redirigé vers le serveur approprié dans le cas d'un service CDN.

Pour calculer ces coûts de transport, nous proposons une approche exploitant les informations de charge des liens réseaux. Bien évidemment, les services applicatifs ne sont pas autorisés à obtenir ce genre d'information, ni n'avoir accès au savoir-faire des opérateurs dont la gestion de la topologie, la politique de routage, le dimensionnement des capacités des liens, etc. Toutefois, notre approche permet d'exploiter ces informations critiques et d'en dissimuler la teneur par respect aux politiques de confidentialité des opérateurs réseaux sans altérer la pertinence des informations délivrées par les matrices de coûts. En effet, ces coûts sont présentés comme des valeurs abstraites qui n'offrent aucune possibilité de déduire des informations sur le réseau, autre que les coûts d'acheminement de bout en bout.

Définition Les coûts de transport d'un chemin réseau est l'accumulation des coûts de transport des liens composants le chemin. Si un des liens constituant le chemin atteint ou dépasse un seuil donné, son coût de transport est fixé à -1 et ainsi sera le coût de transport de tous les chemins réseaux exploitant ce lien.

Exemple Tenons l'exemple illustrée par la figure 2.5. Le coût de transport sur le chemin $\Pi_{4,6}$ entre PID 4 et le PID 6 est la somme $\Sigma(LC_1, LC_5, LC_6, LC_7, LC_8)$. Cependant, en accord avec la politique de sécurité qui tend à éviter l'utilisation des liens réseaux au delà d'un seuil défini par l'opérateur réseau (par exemple 80% de la capacité de chaque lien), les liens dont les conditions de charge ne respectent pas ce critère auront un coût négatif de -1 . Nous considérons que le pourcentage des ressources disponibles restantes (20%) comme marge de sécurité pour le trafic incontrôlable des autres applications.

Ainsi, la première étape pour calculer les coûts de transport est de vérifier si le trafic sur un lien atteint un certain niveau de sa capacité. Dans ce cas, le lien devient temporairement inutilisable. Sur la figure 2.5, $LC_3 = -1$ fait que les chemins $\pi_{2,6}$ et $\pi_{2,7}$ soient inutilisables. Cette politique de sécurité est décidée par chaque opérateur qui en fixe le seuil de façon globale pour tous les liens ou bien séparément pour chacun ou un sous ensemble.

La seconde étape consiste à calculer les coûts de transport de bout en bout entre les PIDs. Si la charge de tous liens respectent la consigne de sécurité, le coût de transport sur un lien i est calculée en utilisant une fonction de coût qui sera présentée dans la section suivante. Le coût $PC_{i,j}$ du chemin entier $\pi_{i,j}$ est par la suite calculé en accumulant les coûts des liens :

$$PC_{i,j} = \sum_{L_u \in \pi_{i,j}} LC_u \quad (2.1)$$

Où L_u est lien réseau numéro u .

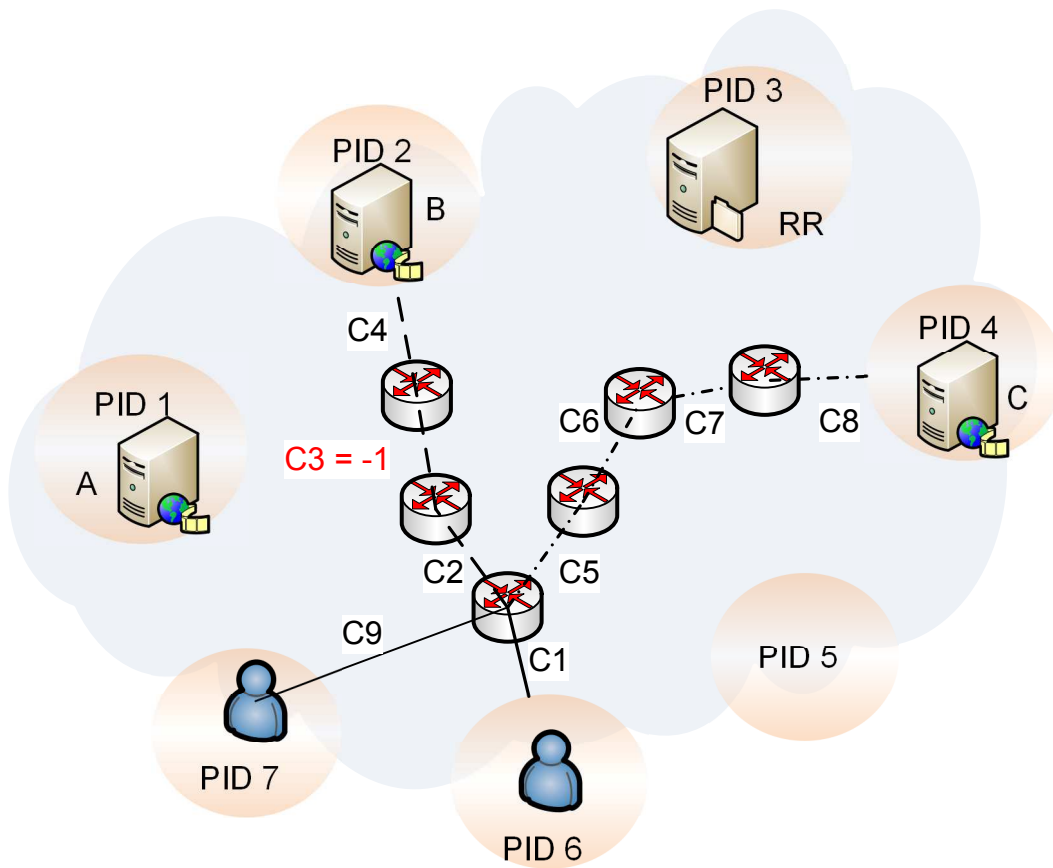


FIGURE 2.5 – Illustration des coûts de transport dans une carte du réseau

2.2.5.2 La fonction de coût

La fonction permettant de calculer les coûts de transport est assujettie à diverses contraintes. Elle doit en effet fournir des informations pertinentes pour les processus de décision des services applicatifs comme la fonction RR des CDNs. D'un autre côté, elle doit permettre de respecter les politiques de protection des informations critiques des opérateurs dans le but d'éviter leur divulgation. En considération de nos critères axés sur l'optimisation de l'utilisation des ressources du réseau et l'amélioration du transport du trafic généré par les services applicatifs collaborants, les métriques d'intérêt dans le calcul du coût de transport sont :

- les taux d'utilisation
- les capacités

des liens réseaux. Le taux d'utilisation d'un lien u du réseau est défini comme le quotient de la charge de trafic LT_u par sa capacité $LCAP_u$. Par ailleurs, il est évident que la fonction de coût devrait croître avec la croissance de ce quotient.

La spécification exacte de cette fonction, en respect des contraintes que nous venons d'énoncer, est propre à chaque opérateur réseau. A charge de chacun de décider de ce que devrait être le coût de transport sur un lien réseau soumis à taux d'utilisation donné. Il existe des approches mathématiques permettant de définir des fonctions personnalisées. Étant donné un ensemble de valeurs exprimant des paires taux d'utilisation, coûts de transport, des techniques d'interpolation comme Newton-Grégoire ou Lagrange ou bien des approximations polynomiales permettent de définir une fonction pour extrapoler le coût d'un taux d'utilisation quelconque. De notre côté, pour le besoin de définir une fonction à utiliser pour nos travaux, nous proposons une fonction exponentielle qui remplit entièrement nos critères avancés dans la section 2.2.5.1.

$$LC_u = \begin{cases} -1 & \text{if } LT_u \geq 0.8 \times LCAP_u \\ a \times \exp\left(b \times \left(\frac{LT_u}{LCAP_u - LT_u}\right)\right) & \text{if } LT_u < 0.8 \times LCAP_u \end{cases} \quad (2.2)$$

Dans cette équation, le seuil limite d'utilisation d'un lien réseau a été fixé à 80%. Rappelons qu'il est fixé par l'opérateur réseau. Le paramètre b est un facteur d'amplification de croissance qui permet de calibrer les coûts d'utilisation suivant des politiques plus ou moins agressives suivant le taux d'utilisation. Le paramètre a est un facteur constant d'amplification et n'a aucune incidence sur la performance du système. Il permet surtout de rendre les valeurs de coûts plus imprévisibles vis-à-vis des fournisseurs de services afin d'éviter tout rapport avec les taux d'utilisation réels des liens. Cette constante peut être modifiée aléatoirement au cours du temps.

Le choix de la fonction de coût s'est porté sur la fonction exponentielle pour sa propriété de croissance, qui est plus forte même que les fonctions polynomiales. Cette

propriété garantit en effet pour un taux d'utilisation élevé, un coût plus élevé que les autres fonctions (à partir d'une certaine valeur bien évidemment). Sachant que notre intérêt se porte plus particulièrement à éviter des situations de forte charge sur les liens du réseau, un coût élevé attribué à ces liens ne serait que pertinent. Dans ce cadre, le terme $LCAP_u - LT_u$ garantit que lorsque la charge de trafic sur le lien s'approche de sa capacité, le coût correspondant tend vers l'infini.

Les figures 2.6 et 2.7 présentent les tracés des fonctions de coût pour des valeurs différentes du paramètre b . On remarque une croissance plus forte des coûts de transport pour des taux d'utilisation plus élevés. Dans l'intervalle $[0 - 50\%]$ la croissance est toute de même plus contenue ce qui fait que l'on considère acceptable des taux d'utilisation dans cet intervalle.

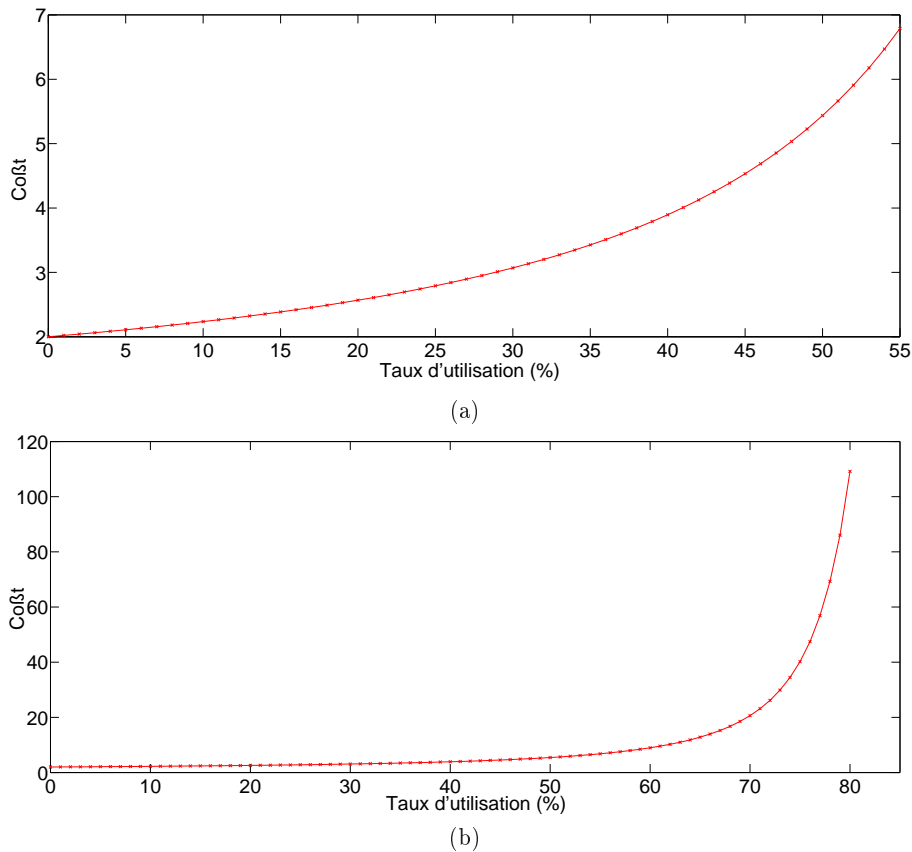


FIGURE 2.6 – Graphe de la fonction de coût avec $a=2$ and $b=1$

2.2.5.3 Analyse des informations exposées par les coûts de transport

Nous avons défini le coût de transport sur un chemin réseau comme la somme des coûts d'utilisation des liens constituant ce chemin. Ainsi, l'information véhiculée par

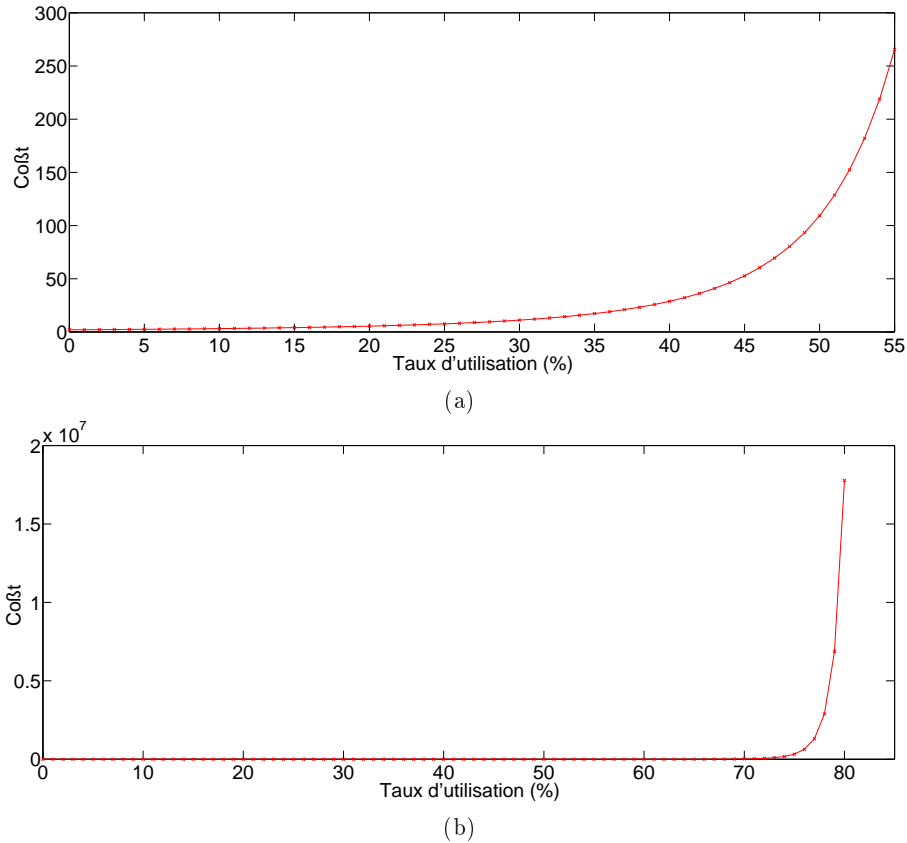


FIGURE 2.7 – Graphe de la fonction de coût $a=2$ and $b=4$

ce coût fournit une indication commune et non séparée sur l'état de chaque lien. La valeur de ces coûts est calculée à partir de la capacité de chaque lien et la charge qui est exercée sur lui par le trafic réseau (et inversement à la bande passante qui lui reste disponible). Contrairement aux autres classes d'informations telles que la distance ou la latence, ces coûts fournissent des indications précises mais relatives sur les conditions de transport :

- précises car calculée à partir des métriques de charge du réseau
- relatives car les chemins seront comparés entre eux et non par rapport à une référence absolue pour éviter la possibilité d'inférer des informations directes sur l'état du réseau

Bien évidemment, l'utilité de cette classe d'information est plus importante pour des services applicatifs avec des contraintes en bande passante et en constance dans le temps comme les services de distribution de contenus ou les transferts d'imposant volume de données. Par ailleurs, les heures de forte charge sur le réseau nécessite aussi une gestion appropriée des ressources quelque soit le service applicatif.

TABLE 2.3 – Comparaison entre les coûts OSPF et les coûts de transport

	OSPF	ALTO
Couche	Couche Réseau	Couche Application
Paramètres de calcul	Capacité	Bande passante et capacité
Algorithm	Dijkstra	Itératif
Paramètres de calcul coût chemin	métriques OSPF, voisins	routes et coût de transport des liens
Objectif	meilleur chemin entre routeurs	meilleur chemin entre PIDs

A un certain degré, cette approche des coûts de transport est sensiblement similaire à celle de certains protocoles de routage comme OSPF [Cis05]. Par défaut, le coût OSPF d'une interface est calculé par la formule

$$coutOSPF = \frac{10^8}{capacite}(bps) \quad (2.3)$$

Notre fonction de coût intègre une variable supplémentaire à la capacité du lien qui est sa bande passante disponible. Mais au final, le coût du chemin est l'accumulation des coûts des liens qui le constituent (voir figure 2.8). OSPF est utilisé par la couche de réseau pour déterminer les meilleurs chemins entre les noeuds du réseau. De façon similaire, en utilisant le protocole ALTO, nous proposons que les opérateurs réseaux fournissent les coûts transport aux services applicatifs afin de leur permettre de déterminer les meilleurs chemins entre les PIDs (voir le tableau 2.3).

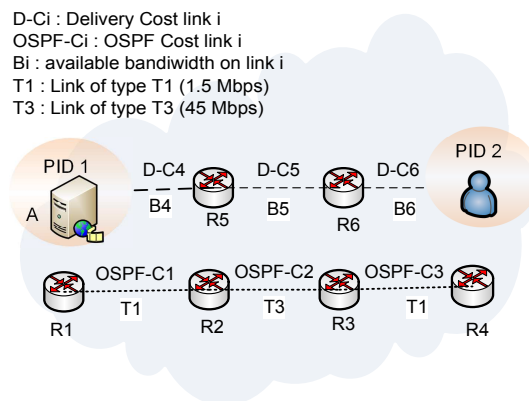


FIGURE 2.8 – Illustration des deux approches de calcul pour les coûts de transport et coûts OSPF

TABLE 2.4 – Comparaison des résultats de différentes fonctions de coût

Chemin	Utilisation des liens (%)			Fonction	Fonction
	L1	L2	L3	Linéaire	Exponentielle
1	20	70	20	11	11.33
2	40	40	40	12	8

L'importance de la fonction de coût Un autre aspect qui revêt d'une importance capitale dans la performance du système est la fonction de coût. Les résultats fournis par cette fonction présentent une interprétation de l'utilisation des liens du réseau. Une fonction linéaire par exemple aboutira à des coûts beaucoup moins élevés qu'une fonction exponentielle ou polynomiale pour des taux d'utilisation moyens ou élevés. Pour illustrer cette différence, prenons l'exemple de la fonction linéaire suivante $L(u) = 10 \times u$, $u \in [0, 1]$, où u est le taux d'utilisation (ici en fraction et non en pourcentage). Nous allons comparer les coûts de transport pour deux chemins différents dans le tableau 2.4 entre la fonction linéaire et la fonction exponentielle que nous avons proposée dans l'équation 2.2 avec $a = 1$ et $b = 2$. Dans ce cas de figure, nous privilégierons le second chemin réseau car l'un des liens dans le premier chemin a atteint 70% de sa capacité alors que les liens dans le second chemin sont tous à 40%. Nous remarquons que les résultats de la fonction linéaire favorise par contre le premier chemin tandis que la fonction exponentielle recommande le second. En effet, le taux d'utilisation élevé du lien $L2$ se traduit en un coût très élevé par la fonction exponentielle qui pénalise donc le premier chemin. En conséquence, les opérateurs réseaux doivent choisir intelligemment la fonction à employer pour le calcul des coûts de transport voire en définir plusieurs pour des situations d'usage différentes.

2.3 L'approche de collaboration active des opérateurs réseaux

L'approche de collaboration active que nous allons présenter est une technique qui met à contribution à la fois les réseaux et les services applicatifs. Elle introduit de nouveaux repères dans les interactions entre les acteurs réseaux et services. Nous développons à travers cette approche l'idée d'un échange mutuel et directe de renseignements entre les acteurs sans faillir aux contraintes de confidentialité, de protection du savoir-faire ou de la sécurité. Nous vous présenterons dans cette section les limites de l'approche de collaboration passive que nous avons introduite dans la section précédente puis nous détaillerons par la suite l'essence de l'approche de collaboration active.

2.3.1 Limites de l'approche passive

L'approche de collaboration passive est, nous rappelons, basée sur un échange d'informations unidirectionnel des réseaux vers les services applicatifs. A cet égard, nous avons retenu le protocole [ALT12] développé par le groupe de travail du même nom à

l'IETF pour permettre aux applications d'obtenir des informations sur les cartes des réseaux et matrices de coûts. Ces données guideront les services applicatifs dans la gestion améliorée du trafic qu'ils génèrent à travers un acheminement sur des chemins réseaux présentant les coûts les plus favorables. Ces interventions se traduisent aux niveau applicatif par des choix de redirection par exemple comme la redirection adéquate des clients CDN vers des noeuds de distribution appropriés [NJWB⁺12], ou bien un choix conséquent d'une liste de pairs pour un utilisateur d'une application P2P, etc. Cependant, deux inconvénients principaux sont soulevés par cette approche.

D'une part, les utilisateurs qui appartiennent au même PID se trouvent en train d'utiliser le même chemin car au final ils sont redirigés vers des noeuds applicatifs appartenant à un autre PID dont le chemin vers le leur présente le coût le plus bas dans la matrice des coûts. A titre d'exemple, la matrice de coûts exposée dans la table 2.5 peut être simplifiée pour donner une table plus compacte (voir table 2.6). Dans cette dernière, on attribut à chaque PID qui contient des clients, un unique PID qui contient les noeuds de distribution qui vont les servir. On voit ainsi dans la table 2.6 que les clients dans le PID1 seront redirigés vers le serveur dans le PID5. Par conséquent, cette approche pourrait conduire à des résultats sous-optimaux. En effet, l'utilisation du chemin réseau concerné va augmenter avec le nombre croissant d'utilisateurs. Dans un contexte de conditions sévères pour le réseau comme les pics de trafic lors les période d'utilisation maximale, un nombre important de redirections statiques entraînant l'utilisation des mêmes chemins réseaux pourrait conduire à des problèmes de congestion, dans un réseau qui pourrait s'auto-défendre avec des mécanismes de perte de paquets. Ces dégradations aboutissent au final sur une médiocre QoE expérimentée par les utilisateurs.

TABLE 2.5 – Un exemple d'une matrice de coûts

PID	1	2	3	4	5
1	0	20	30	22	5
2	15	0	20	25	30
3	30	15	0	20	15
4	20	25	18	0	10
5	15	35	20	10	0

TABLE 2.6 – Un exemple d'une matrice de redirection compacte

PID client	PID serveur
1	5
2	3
3	3
4	5
5	5

D'une autre part, de multiples services applicatifs ayant des besoins et des contraintes différentes les uns par rapport aux autres sont en compétition entre eux pour l'accès aux ressources du réseau. Cependant, dans le contexte d'une collaboration passive, l'échange d'informations est unidirectionnel provenant des réseaux d'opérateurs. Or ceux là mêmes ne sont pas conscients du contexte (besoins, contraintes, types, étendues, etc) des différentes applications qui sont entrain d'utiliser leurs ressources et offrent ainsi les mêmes informations menant aux mêmes décisions aux niveaux des différents services applicatifs. En l'occurrence, toute tentative de rationnement ou personnalisation serait inefficace sans connaissance préalable des bénéficiaires. En conséquence, il peut en résulter des mauvaises performances dans le cas d'une forte croissance du trafic par exemple entraînant une sur-utilisation de certains liens. Afin d'éviter de telles situations, nous proposons une nouvelle approche, où les différents acteurs s'expriment et collaborent activement.

2.3.2 Proposition d'amélioration de l'approche passive

Afin de remédier aux limites de l'approche passive exposées précédemment, nous proposons une évolution du protocole de communication ALTO et l'intégration de nouvelles classes d'informations dont la matrice de coûts de proportions, que nous appellerons désormais matrice de proportions, et la matrice de contraintes. Ces nouvelles matrices permettront d'avoir un échange mutuel de renseignements entre les acteurs réseaux et services. A ce titre, ces évolutions sont introduites par la nouvelle interface CINA, Collaboration interface between Networks and Applications (CINA), qui intègrent des services additionnels notamment la possibilité pour les services applicatifs de bénéficier de certains services réseaux exclusifs comme le multicast.

2.3.2.1 La matrice de proportions

Nous introduisons une nouvelle classe d'informations dans cette matrice de coûts décrivant non plus des coûts d'acheminement comme précédemment, mais des proportions de trafic ou de nombre de sessions à acheminer entre les différents PIDs. Les gestionnaires de services applicatifs exploitent ces informations pour réaliser un aiguillage précis du trafic de leurs entités applicatives afin de se conformer aux proportions qui leurs ont été communiquées par les opérateurs réseaux. Un exemple d'une matrice de coûts fournissant des informations sur les proportions du trafic entre les différents PIDs contenant des utilisateurs (PIDs 1, 2 et 4) et des serveurs (PIDs 3 et 5) est exposée dans la figure 2.9. Dans cet exemple, où le type de coût est pourcentage, 70% du trafic reçu par les clients du service appartenant au PID 1 devrait être livré par des entités (noeuds de distribution, data-centres, pairs, etc) dans le PID 3, le reste par des entités dans le PID 5.

Cette nouvelle classe d'informations que nous proposons permet de contrer efficacement les limites des coûts de routage statiques précédemment introduits. En effet, définir des proportions de trafic à respecter assure une bonne répartition de la charge globale sur les différents chemins et en conséquent sur les différents liens du réseau. Il s'agit

bien évidemment de prendre en compte l'état d'utilisation des liens dans le calcul de ces proportions. Les résultats escomptés de l'emploi de cette approche sont dans l'intérêt de tous les acteurs. Dans les réseaux, la charge de trafic sera répartie de façon optimale ce qui évitera ou bien limitera les situations de surcharge pour certains liens du réseau tandis que d'autres sont sous exploités. Les services applicatifs de leur côté bénéficieront d'une meilleure conditions de transport grâce à une meilleure disponibilité des ressources réseaux d'où des risques réduits de congestion et perte de paquets. Enfin, les utilisateurs eux-mêmes témoigneront d'une meilleure QoE assurée par un acheminement performant de leurs données.

2.3.2.2 La matrice de contraintes

Les opérateurs réseaux sont responsables de la mise à disposition des services applicatifs de cette nouvelle classe d'informations. Mais afin d'être en mesure de déterminer les proportions de trafic ou de nombre de sessions à suivre par les différents services applicatifs, les opérateurs réseaux doivent maîtriser un certain nombre d'informations relatives aux services applicatifs qu'ils souhaitent collaborer avec afin d'apporter des informations optimales et personnalisées à chaque application. Nous avons identifié deux classes d'informations nécessaires et suffisantes pour atteindre l'objectif principal qui est l'optimisation de l'acheminement du trafic dans le réseau. Dans ce contexte particulier, l'opérateur réseau a besoin de connaître :

- La localisation et la capacité des entités applicatives
- Les prévisions de trafic du service dans le réseau

Conditions nécessaires En effet, il est nécessaire de savoir où se trouvent les entités serveurs d'un service particulier sans quoi, l'opérateur ne peut rien calculer. En marge, ignorer la capacité ou la disponibilité de ressources pour ces entités risque d'aboutir à des proportions de trafic qu'elles ne peuvent assumer. A titre d'exemple, un opérateur réseau ne peut pas recommander à un opérateur CDN de rediriger un certain nombre de ces clients vers un noeuds de distribution dépassant la capacité de ce dernier. De façon similaire, il est nécessaire de connaître les quantités de trafic à devoir gérer dans

```
{
  "cost-mode" : "percentage",
  "cost-type" : "proportional",
  "map-vtag" : "225006996",
  "map" : {
    "PID1": {"PID3":70, "PID5":30},
    "PID2": {"PID3":65, "PID5":35},
    "PID4": {"PID3":100}
  }
}
```

FIGURE 2.9 – Illustration d'une matrice de proportions

le réseau afin de pouvoir les répartir de façon optimale.

Conditions suffisantes Par ailleurs, ces informations sont suffisantes de la part des gestionnaires de services applicatifs pour résoudre le problème d'optimisation. En effet, en connaissant la localisation des entités sources de données et leurs capacités, nous savons où et quel nombre de clients rediriger vers ces entités ou bien quelle quantité de trafic desservir à partir de ces entités. L'autre donnée relative aux prévisions de trafic permet de savoir quelle quantité de trafic ou quel nombre de client s'attendre à avoir dans chaque PID d'où la possibilité de déterminer les proportions entre chaque PID contenant des clients et les différents PIDs contenant des serveurs. il ne peut pas recommander l'équilibrage du trafic pour un service Cloud sans la connaissance de l'ensemble du trafic entrant et / ou sortant requis pour chaque PID. Pour permettre de telles interactions entre les couches, nous introduisons un nouveau service, le Service de cartographie de contraintes, ce qui permet CCSP pour exposer des informations à NOs sans divulguer des informations essentielles telles que les adresses IP.

Cette matrice de contraintes qui est fournie par les gestionnaires de services exposent les informations requises par les opérateurs réseaux pour achever le calcul des valeurs optimales de proportions. Elle est fidèle au format de structure introduit par le protocole ALTO. Les informations exposées sont relatives aux PIDs définis par la carte du réseau. La matrice de contraintes exposent deux classes d'informations dans la même structure profitant des services additionnels rajoutés à ALTO comme les matrices à plusieurs coûts. Cette matrice expose au final deux contraintes (voir figure 2.10). La première contrainte décrit l'emplacement et les capacités des entités serveurs. A titre d'exemple, dans la figure 2.10, cette capacité est exprimée en terme de capacité à servir une quantité de trafic en Gbps d'où le serveur dans le PID 3 est capable de gérer *20 Gbps* de trafic. La deuxième contrainte expose les prévisions par le gestionnaire de service du trafic vers chaque PID. Par exemple, 6 Gbps de trafic sont prévus pour des clients dans le PID 1. Pour déterminer ces valeurs, les gestionnaires de services peuvent s'appuyer sur des outils d'analyse statistique de leurs historiques de services.

2.3.3 L'échange bidirectionnel d'informations

Les opérateurs réseaux et gestionnaires de services utilisent L'interface CINA pour échanger les différentes cartes et matrices d'informations nécessaires à l'exécution des procédures de calculs et d'optimisations. De cette façon, les opérateurs réseaux sont capables de répondre aux sollicitations et contraintes des services applicatifs tout en assurant une utilisation optimale de leurs ressources réseau. Cet objectif engage une réponse personnalisée à chaque service applicatif. Le diagramme de flux qui décrit la procédure de collaboration simplifiée pour l'échange d'informations entre réseaux et applications est représenté dans la figure 2.11.

2.3.4 Le problème d'optimisation

Les nouveaux services d'échange d'informations que nous avons introduits dans l'interface CINA ont pour objectif de permettre aux acteurs de la collaboration de remédier aux inconvénients de la collaboration passive précédemment présentée. Dans ce contexte d'interaction inter-couches, nous proposons l'utilisation des information échangées dans la mise en oeuvre de procédures de calculs et d'optimisation au niveau du réseau. En effet, le réseau constitue l'environnement commun aux différents acteurs. L'accès aux ressources disponibles doit être géré de façon optimale entre les différents services applicatifs en fonction des besoins et des contraintes de chacun. En conséquence, pour des

```

{
  "constraint-mode" : ["Gbps","Gbps"],
  "constraint-type" : ["surrogate-
capacity","inbound traffic"],
  "map-vtag" : "42550001",
  "map" : {
    "PID1" : [{"0","6"}],
    "PID2" : [{"0","4"}],
    "PID3" : [{"20","6"}],
    "PID4" : [{"0","10"}],
    "PID5" : [{"5","0"}],
  }
}
    
```

FIGURE 2.10 – Illustration d'une matrice de contraintes

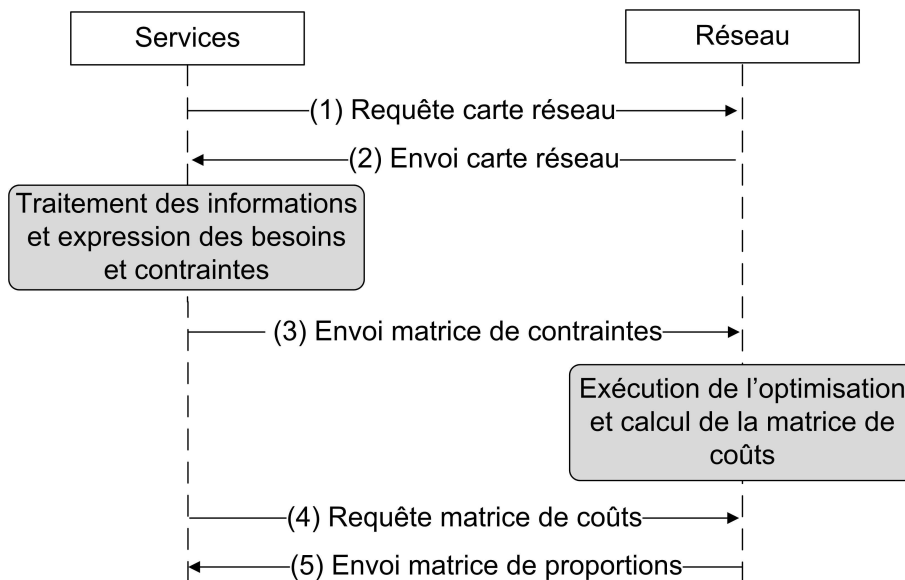


FIGURE 2.11 – Diagramme de flux de la procédure d'échange d'informations utilisant l'interface CINA entre opérateurs réseaux et gestionnaires de services applicatifs

raisons évidentes, il apparaît que l'opérateur réseau est le plus à même de fédérer l'ensemble des besoins et contraintes à la fois de ses ressources et services mais aussi celles des divers services applicatifs qui utilisent son réseau. Cette mission lui incombe de définir la stratégie d'optimisation à suivre. Nous présentons l'approche de modélisation mathématique du problème et détaillons sa résolution dans les sections suivantes.

2.3.4.1 Modélisation mathématique du problème

Considérons un réseau de transport qui a été étudié et modélisé par une topologie constituée de N PIDs relié entre eux par M liens réseau, qu'on note $L_u, u \in [1, M]$. Chaque lien u est caractérisé par sa capacité de transport qu'on dénote $LCAP_u$. Soit $S, S = S_i$, l'ensemble des entités applicatives source de données ou serveurs, où i dénote le PID auquel appartient le serveur S_i , ç-à-d sa localisation dans la topologie, et $SCAP_i$ sa capacité. Nous définissons la variable $X = x_{ij}$ comme la proportion de trafic entre les utilisateurs dans le PID i and les serveurs dans le PID j . Cette variable x_{ij} peut aussi modéliser un nombre de sessions au lieu de proportions de trafic. Dans le cas échéant, nous passerions d'un modèle plus générique qui repose sur la quantité de données à un modèle plus restreint qui repose sur un nombre de sessions.

Nous introduisons une marge de sécurité qui garantit un certain seuil, par exemple 80% de l'utilisation des liens réseau. Cette limite est fixée par les opérateurs réseaux et est personnalisable pour chaque lien du réseau séparément. Son utilité est capitale pour faire face à une augmentation incontrôlable du trafic. Dans ce cas, tout lien dont le taux d'utilisation dépasse les 80% de sa capacité, présente désormais un coût négatif égal à -1 , ce qui se traduit par le fait que le lien est temporairement inutilisable. Par contre, si le trafic dans le lien est en dessous ce ce seuil, alors la fonction de coût (équation 2.2) est utilisée pour calculer le coût d'utilisation LC_u du lien L_u fonction du trafic réseau sur le lien LT_u et sa capacité $LCAP_u$. Nous rappelons que le coût d'utilisation de chemin réseau π_{ij} entre les PIDs i et j est la somme des coûts des liens composant le chemin comme le définit la fonction 2.1.

Matrice d'utilisation des liens $LU = [LU_u^{ij}]$ est la matrice qui définit les taux d'utilisation des liens réseau en fonction des proportions de trafic x_{ij} routées vers les différents liens. La matrice a une dimension $M \times K$ où M est le nombre de liens et K est le nombre de variables $X = x_{ij}$. Les vecteurs $[LU_1^{ij}, LU_2^{ij}, \dots, LU_M^{ij}]$ qui représentent les colonnes de la matrice LU déterminent les liens qui sont utilisés par chaque variable x_{ij} . Les vecteurs $[LU_u^1, LU_u^2, \dots, LU_u^K]$ qui représentent les lignes de la matrice indiquent l'ensemble des variables x_{ij} qui utilisent le lien L_u . A titre illustratif, considérons la matrice présentée dans la table 2.7. Dans cet exemple, $LU[1, 3] = 1$ indique que lien L_1 subit une proportion de trafic définie par la variable x_3 . L'ensemble des paramètres du modèle mathématique sont présentés dans la table 2.8.

2.3.4.2 Définition du problème

Le problème d'optimisation que nous allons présenter a pour objectif de déterminer les valeurs optimales des proportions de trafic entre les différents PIDs qui permettront d'atteindre un équilibre satisfaisant entre l'amplitude de la charge de trafic exercée sur les différents liens de la topologie réseau et la qualité de service attendue par les services applicatifs. Pour répondre à ces besoins, nous avons intégré deux axes d'optimisation dans le problème

Optimisation des coûts de transport Le premier objectif que nous proposons d'atteindre est d'optimiser le coût global d'utilisation du réseau en répartissant la charge de trafic de façon intelligente sur les liens réseaux. Nous définissons le coût d'utilisation du réseau comme la somme des coûts de tous les liens (voir équation 2.4).

$$NetCost = \sum_{u=1}^M LC_u \quad (2.4)$$

Nous cherchons à minimiser le coût global $NetCost$ en déterminant les valeurs optimales pour x_{ij} . Nous allons donc spécifier la relation entre ces paramètres. La nouvelle charge de trafic exercée sur le lien L_u considérant l'apport des proportions x_{ij} est calculée dans l'équation 2.5.

$$LT_u^{opt} = LT_u + LU_u^{ij} \times X \quad (2.5)$$

$$LU_u^{ij} \times X = \sum_{ij \Delta u} x_{ij} \quad (2.6)$$

où $ij \Delta u$ inclut tout le trafic qui va utiliser le lien L_u . En effet, le vecteur binaire LU_u^{ij} spécifie l'ensemble des variables x_{ij} qui vont utiliser le lien L_u . Ainsi nous pouvons définir la relation entre le coût global $NetCost$ and le vecteur proportions X dans l'équation 2.7.

$$NetCost = \sum_{u=1}^M a \times \exp \left(b \times \frac{LT_u + LU_u^{ij} \times X}{LCAP_u - (LT_u + LU_u^{ij})} \right) \quad (2.7)$$

Optimisation des chemins réseaux L'optimisation des coûts de transport peut être accompagnée par une optimisation du nombre de liens utilisés. L'utilisation d'entités serveurs plus proches des utilisateurs finaux présente des avantages indéniables quand cette opportunité est en adéquation avec les disponibilités de réseau de transport. Nous définissons à cet égard un paramètre d'utilisation des liens réseaux, qui indique le nombre global de liens utilisés par les services applicatifs (voir équation).

$$NetLink = \sum_{k=1}^K \sum_{u=1}^M LU_k^u \quad (2.8)$$

Le vecteur binaire LU_k^u , *colonne de la matrice LU* ($M \times K$) fournit le nombre de liens utilisés par chaque variable x_{ij} , traduite en variable x_k pour les besoins du calcul matriciel.

Le problème d'optimisation global En intégrant les deux axes d'optimisation dans notre approche, nous définissons un nouveau problème d'optimisation multi-objectif appelé aussi optimisation vectorielle. Ce problème d'optimisation peut être défini comme suit :

$$\begin{aligned} \text{Minimize} \quad & F(x) = [F_1(x), F_2(x), \dots, F_p(x)] \\ \text{subject to} \quad & g_i(x) \leq 0, i = 1, 2, \dots, n \\ & h_j(x) = 0, j = 1, 2, \dots, m \end{aligned} \tag{2.9}$$

C'est un problème sous contraintes d'égalités h_j et inégalités g_i . Il existe des techniques sophistiquées pour résoudre ce type de problèmes et qui peuvent être basées sur des pondérations des objectifs ou bien des algorithmes révolutionnaires ou génétiques par exemple. Les auteurs de cet article [MA04] exposent un panel de techniques capables de fournir des solutions à ce genre de problèmes et y détaillent les approches.

Dans notre approche, certaines considérations sont influentes sur le choix de la technique à utiliser. D'abord la dimension des fonctions sujet à optimisation de notre problème peuvent avancer un grand nombre de variables atteignant facilement la centaine. Il y a aussi le caractère non linéaire de la fonction de coûts qui rajoutent une complexité de calcul. Pour garantir des performances acceptables avec un temps de calcul contenu, nous avons décidé d'orienter la technique de résolution vers la méthode de la somme pondérée des fonctions objectives.

$$F = \sum_{i=1}^p \omega_i F_i(x) \tag{2.10}$$

A condition que tous les poids ω_i de l'équation 2.10 soit positifs, cette approche assure que la solution du problème est "*Pareto Optimal*" [Zad63], c-à-d minimiser l'expression dans l'équation 2.10 est suffisant pour atteindre l'optimalité *Pareto*.

Nous avons défini en conséquence une nouvelle fonction multi-objectif *Net* que nous cherchons à minimiser (équation 2.11).

$$Net = (\alpha \times NetCost) + (\beta \times NetLink) \tag{2.11}$$

α et β sont les poids respectifs de chaque objectif. le choix des valeurs de ces pondérations traduit la préférence de l'opérateur réseau envers l'un ou l'autre des axes d'optimisations. Le problème global d'optimisation qui concerne notre système est donc défini ainsi :

$$\begin{aligned}
 & \text{Minimiser} && \text{Net}(X) \\
 & \text{sujet à} && LT_u + LU_u^{ij} \times X \leq LCAP_u \quad \forall u \\
 & && \sum_i x_{ij} \leq SCAP_j \quad \forall i, j \\
 & && \sum_j x_{ij} = T_i \quad \forall i, j \\
 & && x_{ij} \geq 0 \quad \forall i, j
 \end{aligned} \tag{2.12}$$

La première contrainte est une contrainte d'inégalité qui indique que la charge de trafic sur chaque lien, qui est la somme de la charge initiale et la charge rajoutée par les variables x_{ij} ne doit pas dépasser le seuil décidé par l'opérateur (par exemple 80% de la capacité du lien). La deuxième contrainte spécifie que le total du trafic desservis par chaque entité du service applicatif ne doit pas dépasser sa capacité. La contrainte d'égalité garantit que le trafic total destiné à chaque PID est égal aux besoins exprimés par le service applicatif à travers la matrice des contraintes. La dernière contrainte stipule que les variables x_{ij} doivent être toujours positives.

2.3.4.3 La résolution du problème

Le problème (2.12) est un problème d'optimisation non linéaire sous contraintes d'égalité et inégalité linéaires caractérisé par une imposante dimension de variables. Le premier défi consiste à déterminer l'existence d'une région faisable \mathfrak{R} de recherche (équation 2.13), c'est à dire vérifier si l'ensemble des conditions découlant des contraintes exposées par le réseau (par exemple les capacités des liens) et des contraintes exposées par les services applicatifs (comme les prévisions de trafic et les capacités des serveurs) peuvent être satisfaites.

$$\mathfrak{R} = \{x \in R^n \mid h(x) = 0, g(x) \leq 0\} \tag{2.13}$$

Le second défi consiste à déterminer l'existence d'une solution optimale dans cette région. Le théorème Karush-Kuhn-Tucker (KKT) [Kuh82] fournit un ensemble de conditions nécessaires et suffisantes pour garantir l'existence d'une solution optimale qui tient sous les contraintes exposées. Toutefois, la dimension du problème et des contraintes exigent des ressources de calcul considérables pour une approche classique. En conséquence, nous avons employé un algorithme que nous présenterons un peu plus loin et qui est capable de gérer de larges problèmes d'optimisation.

Préparation de la résolution L'optimisation que nous cherchons à atteindre est née d'une problématique que nous avons identifiée et exposée dans les sections précédentes. Sa transposition en problème mathématique pur nous a permis de concrétiser nos objectifs. A travers notre modèle, nous cherchons à déterminer les quantités de trafic optimales entre les clients et les serveurs qui sont localisés dans différents PIDs du réseau. Nous avons traduit ces quantités en proportions de trafic sur chaque lien. Ces

variables nous sont inconnues et nous cherchons à utiliser un algorithme d'optimisation pour les déterminer. Cet algorithme devra satisfaire des contraintes que nous lui spécifieront. Ces contraintes vont réduire l'espace des solutions. L'algorithme devra trouver une solution optimale dans l'espace disponible quand c'est possible. En effet, il arrive aussi que les contraintes soient tout simplement impossible à satisfaire. concrètement dans un réseau de transport, il pourrait s'agir d'une quantité de trafic supérieure à la capacité du réseau. Ces contraintes proviendront d'un côté des capacités de chaque lien du réseau. Ainsi l'algorithme devra trouver des proportions de trafic, qui en s'ajoutant au trafic existant ne dépasseraient pas une certaine limite de capacité. En conséquence, ces capacités sont fournies à l'algorithme. Les contraintes proviennent aussi des besoins des applications. Les clients s'attendent à recevoir leurs flux de données. En conséquence, l'algorithme doit être en mesure de fournir des valeurs qui au final correspondent à la quantité de trafic attendue dans chaque PID.

La préparation de la résolution consiste donc en la transposition des paramètres du modèle en variables d'entrée et de sortie pour l'algorithme afin de définir les contraintes et déterminer la solution optimale.

L'algorithme utilisé Il s'agit de l'algorithme programmation séquentielle quadratique Sequential Quadratic Programming (SQP) [BT95, BTK94, Bar08]. SQP apporte une méthode très efficace pour la résolution d'imposants problèmes d'optimisation. Nous allons fournir une brève présentation de l'approche de résolution de problèmes employée par l'algorithme et ses avantages sans entrer dans les détails. Des exposés bien détaillés sur la méthode SQP sont disponibles [BT95, BTK94, Bar08].

L'idée de base de SQP est de modéliser le problème d'optimisation sous la forme d'un sous problème de programmation quadratique à une certaine solution approximative x^k . Ensuite, l'algorithme utilise cette solution pour trouver une meilleure approximation x^{k+1} . L'itération de cette procédure permet de construire une suite d'approximations qui pourrait converger vers une solution x^* . Cette approche peut être assimilée à une extension des méthodes Newton et quasi-Newton dans la résolution des problèmes d'optimisation sous contraintes. SQP est aussi caractérisé par le fait qu'il n'appartient pas à la famille des méthodes de recherche stricte dans la région faisable ("*feasible-point methods*"). Ces techniques exigent que la suite de solutions déterminées par la séquence d'itération se trouve dans la région faisable. SQP est plus tolérant dans le sens où les solutions peuvent se trouver à l'extérieur de la région faisable. Cette tolérance apporte un gain considérable en performance car déterminer un point dans la région faisable peut être aussi difficile que la résolution même du problème, surtout en présence de contraintes non linéaires. Un autre intérêt en faveur de l'emploi de la méthode SQP est l'existence d'algorithmes reconnus rapides et efficaces dans la résolution des problèmes de programmation quadratique. En considération de ces aspects et des évaluations de performances de l'algorithmes SQP [Sch86], l'approche fournit des résultats intéressants pour les problèmes de grande envergure qui nous concerne en termes d'efficacité et de précision.

Chapitre 3

Application des approches de collaboration

Nous avons présenté dans le précédent chapitre les approches de collaboration que nous proposons de mettre en application pour améliorer les performances des réseaux et des services. Nous avons détaillé l'architecture du système, les modalités d'exploitation et les procédures d'optimisation. Dans ce chapitre, nous présenterons des cas d'application de l'approche à des services de distribution de contenus fournis par CDN ou bien par application P2P. Nous discuterons des différents niveaux de collaboration et avancerons des propositions qui permettent d'améliorer l'utilité des approches de collaboration.

3.1 Application de l'approche aux services CDN

Dans cette section nous présenterons l'intégration des approches de collaboration dans les architectures de type CDN et leur application dans pour les services qu'elles fournissent. Nous présenterons brièvement les principaux modules fonctionnels d'un architecture CDN puis nous détaillerons les interactions nécessaires pour l'exploitation de nos propositions.

3.1.1 Présentation d'une plateforme CDN

Une plateforme d'un réseau de distribution de contenus peut revêtir d'un niveau de complexité considérable suivant la taille et les services fournis par le CDN. Néanmoins, certains systèmes sont de-facto communs à tous les CDNs mais se différencient par leurs niveau de complexité. Ces systèmes qui assurent les fonctions de base incluent :

- Système de contrôle : ce système assure des fonctions vitales comme le contrôle des autres systèmes du CDN ou les interactions avec les acteurs externes.
- Système de logging : ce système est responsable de la collecte d'informations sur le fonctionnement des services CDN. Ces informations peuvent viser les performances des services ou bien servir à la facturation des services rendus par le CDN

par exemple.

- Système de routage et distribution : ce système assure le routage à la fois des données et meta-données relatives à l'intérieur du réseau et des utilisateurs desservis par le CDN.
- Système de transport : ce système est responsable du transport des données.

Nous proposons d'étudier l'intégration de notre proposition d'approche collaborative dans la plateforme CDN d'Akamai [NSS10], le leader mondial mondial des réseaux de distribution de contenus.

3.1.1.1 La plateforme d'Akamai

Nous présenterons dans cette section les principales composantes de la plateforme d'Akamai 3.1 afin d'analyser par la suite les implications de l'application de notre proposition de collaboration avec les réseaux de transport sur ce système.

les serveurs Edge Ces serveurs constituent les noeuds de distribution qui desservent les clients finaux en contenus demandés. Ils implémentent divers fonctionnalités comme la gestion des chemins vers des contenus statiques ou dynamiques, l'authentification, etc.

Le système de Mapping Ce système gère le routage du trafic du CDN. Il se base sur des informations temps réel collectées depuis le réseau de distribution et des statistiques collectées du historique de fonctionnement afin de décider du routage du trafic dans le réseau du CDN. Ce système se scinde en deux modules :

- Le Macro Mapping : Appelé "*Mapping Scoring*" by Akamai, il est responsable de la création du carte topologique qui capture l'ensemble l'état des liaisons dans tout l'univers Internet. Il s'agit plus précisément de segmenter le réseau mondial en régions et de déterminer la bonne santé des liaisons entre ces régions. Cette procédure se base sur un ensemble d'informations collectées depuis des pings, traceroute, informations BGP, logs, etc.
- Le Mapping temps-réel : ce module est responsable de la création des matrices de redirection utilisée par les fonctions de "*Request Routing*" pour rediriger les utilisateurs vers le meilleur noeud de distribution.

Le système de communication et de contrôle Ce système est responsable du contrôle de l'ensemble du parc de machines du CDN (plus de 60,000 serveurs) à des fins de monitoring, de configuration, de mise à jour, etc.

Le Système de collecte et analyse des données Ce système utilise divers mécanismes de collecte d'informations, incluant :

- Les logs : ces logs contiennent des informations sur l'utilisation des ressources du CDN qui permettent de réaliser par exemple des analyses de performances ou d'audit ou bien de facturer les services rendus par le CDN.

- Monitoring temps réel : à travers un système distribué de base de données, Akamai est capable d'accéder en quasi temps réel aux informations relatives aux performances de chaque composant software ou hardware du CDN
- Analyse et reporting : ce module permet à la fois à Akami et ces clients de visualiser des informations de performance et de trafic qui concernent leurs services

L'Agent de Monitoring Cet agent surveille les performance des services et des réseaux de transport. Ils sont capables de réaliser des mesures de pings, d'exécuter des procédures traceroute, etc.

Le Portail de gestion Il s'agit d'un portail web qui offre divers fonctionnalités de configuration aux clients du CDN pour leur permettre de contrôler et visualiser les contenus qu'ils souhaitent distribuer à travers le CDN.

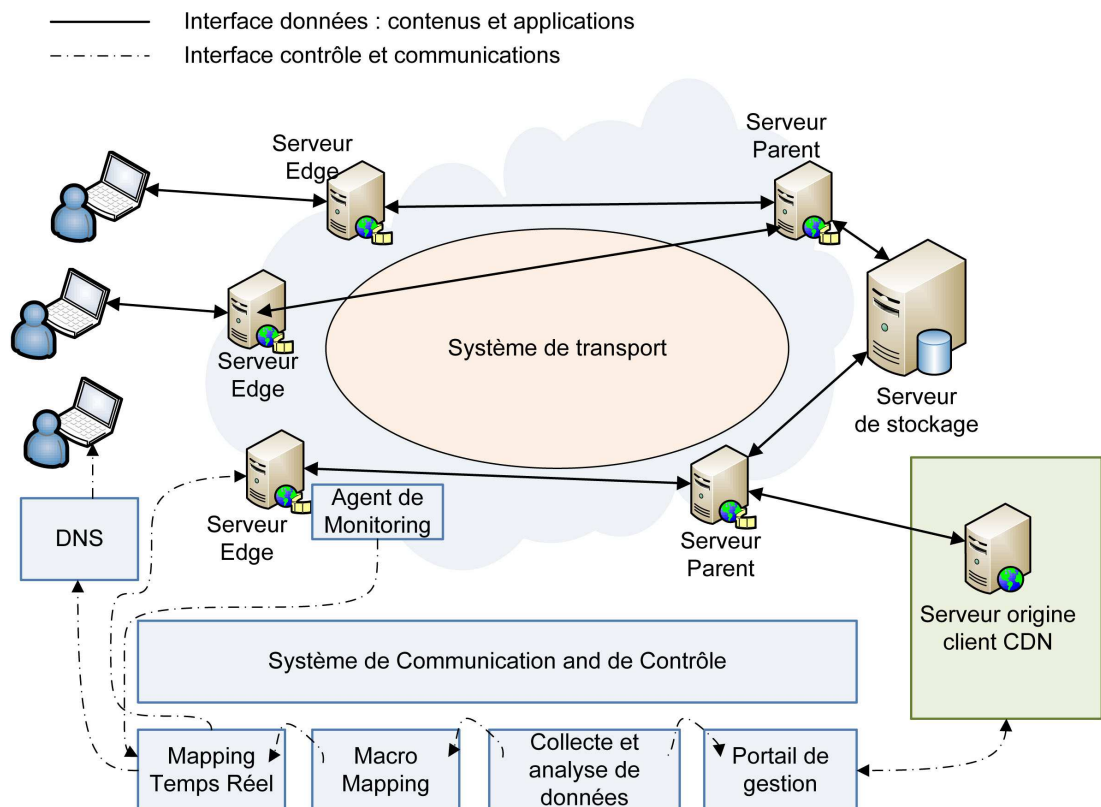


FIGURE 3.1 – Illustration de la plateforme CDN d'Akamai

3.1.2 Intégration des modules de collaboration dans une plateforme CDN

L'application de l'approche de collaboration entre un opérateur CDN et un opérateur réseau nécessite l'intégration du système de collaboration CINA présenté dans la section 2.1.2.1 dans les plateformes des deux acteurs. Nous présentons dans cette section les détails de l'intégration de l'architecture CINA (voir figure 2.1) dans la plateforme CDN d'akamai (voir figure 3.2).

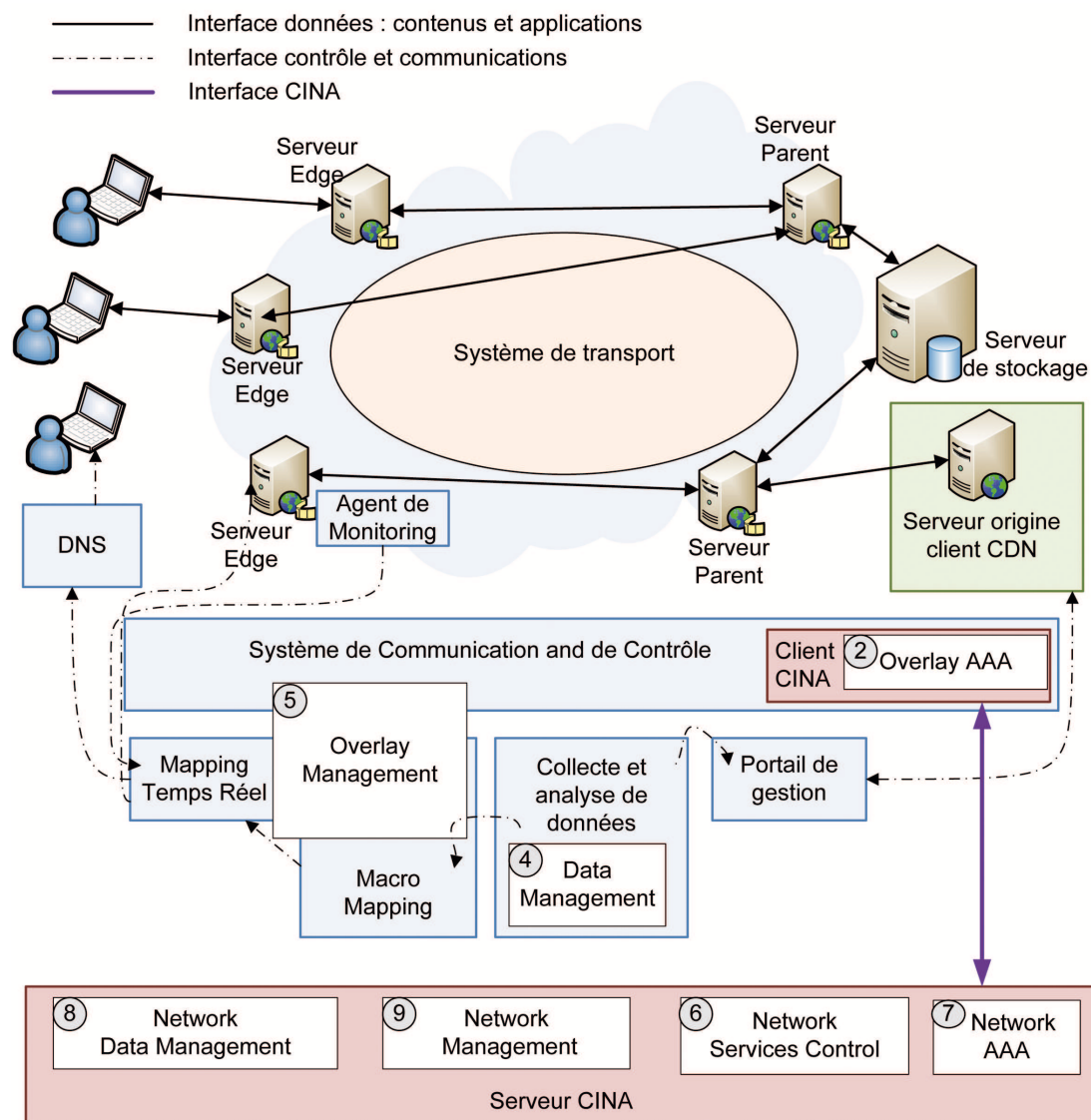


FIGURE 3.2 – Illustration de l'intégration de l'architecture CINA dans la plateforme CDN d'akamai

L'opérateur CDN se doit d'implémenter le client CINA qui communiquera avec le serveur CINA implémenté par l'opérateur réseau. Ce client est intégré au module de communication et de contrôle qui est responsable des interactions avec les acteurs externes notamment les opérateurs réseaux. Les échanges sont opérés par le module fonctionnel "*Overlay Management*" de l'architecture CINA. Ce module qui est responsable aussi du traitement des informations et de la gestion du routage du trafic applicatif du CDN. Il dispose de fonctions implémentées dans les systèmes *Macro Mapping* et *Mapping Temps Réel* de la plateforme Akamai. Ces fonctions assurent l'intégration des recommandations fournies par les opérateurs réseaux dans la politique de sélection de noeuds de distribution du CDN. La collecte et la sauvegarde des informations récupérées sont gérées par le module fonctionnel "*Data Management*" implémenté dans le système Akamai de collecte et d'analyse de données. Ce module fournit en outre les informations nécessaires à la préparation des contraintes exposées par le CDN, à savoir ses prévisions de trafic et ses capacités de services.

Le serveur CINA est déployé par l'opérateur réseau. La fonction de ce système est passive vis-à-vis du réseau, ç-à-d, il n'induit pas de modifications dans la politique de fonctionnement du réseau. Le serveur est en intégration avec les systèmes de monitoring du réseau qui lui fournissent les métriques nécessaires aux procédures de calculs et d'optimisation.

Cette intégration souligne surtout la capacité du système de collaboration CINA d'agir en mode passif, ç-à-d une collaboration unidirectionnelle de la part des opérateurs réseaux, ou bien en mode actif avec une collaboration mutuelle des deux acteurs dans le contexte de services CDN. Dans ce cadre, il s'agit avant tout de définir le périmètre de la collaboration entre opérateurs CDN et réseau en fonction de leurs politiques et besoins.

3.2 Application de l'approche aux applications P2P

Malgré un notable déclin ces dernières années de l'utilisation des applications P2P, elles présentent toujours des avantages indéniables inhérents à leur architecture distribuée et continuent de susciter l'intérêt d'une utilisation parallèle aux infrastructures client-serveur type CDN pour un allègement de charge. Nous évoquerons dans cette section les opportunités de l'application des approches de collaboration que nous proposons pour les applications P2P. Nous présenterons par la suite des travaux que nous avons menés pour la conception d'une architecture optimisée pour les applications P2P *streaming*.

3.2.1 Intégration dans les architectures P2P existantes

Le groupe de travail ALTO à l'IETF a réalisé des travaux d'investigations relatifs à l'intégration de l'architecture ALTO, noyau du système CINA, dans les applications P2P [ALT12]. Ces travaux ont abouti à la spécification de deux voies d'intégration possibles dans les systèmes P2P.

Intégration dans le tracker Plusieurs applications P2P sont gérées par des trackers qui assurent les opérations de sélection de pairs auxquels sera connecté chaque client P2P. En effet, le tracker dispose de la liste de tous les clients connectés au réseau P2P. Pour chaque client, il sélectionne une liste de pairs avec qui il échangera les données. Cette sélection peut être optimisée à travers des informations relatives aux conditions de transport dans le réseau de l'opérateur. Pour avoir accès à ces informations, le tracker implémente un client CINA qui lui permet de communiquer avec le serveur CINA de l'opérateur (voir figure 3.3). Il a besoin par ailleurs d'implémenter les fonctionnalités de traitement et de sauvegarde des informations collectées.

Intégration dans le client P2P Dans ce mode d'intégration, c'est le client P2P qui est responsable de collecter les informations depuis le serveur CINA de l'opérateur réseau (voir figure 3.4). En partant d'une liste de pairs fournie par une entité centralisée comme un tracker ou bien distribuée comme les tables de hachage distribuée "*Distributed Hash Table*" (DHT), il sélectionne les meilleurs peers relativement aux recommandations de l'opérateur réseau. Cette technique permet aux utilisateurs des applications P2P totalement distribuée de profiter de la collaboration d'un opérateur réseau au sur-coût d'une charge plus importante sur son serveur. En effet, chaque client devra récupérer les informations depuis le serveur de l'opérateur et les mettre à jour au cours de son utilisation.

Ces voies d'intégration permettent aux applications P2P de profiter des informations que peuvent leur fournir les opérateurs réseaux afin d'améliorer le processus de sélection de pairs et en contre partie améliorer l'utilisation des ressources réseaux. Il s'agit d'une application de l'approche de collaboration passive où toute l'optimisation est réalisée au niveau applicatif. L'application de l'approche de collaboration active pour ce type d'architecture paraît impossible. D'un côté, l'intégration du système de collaboration

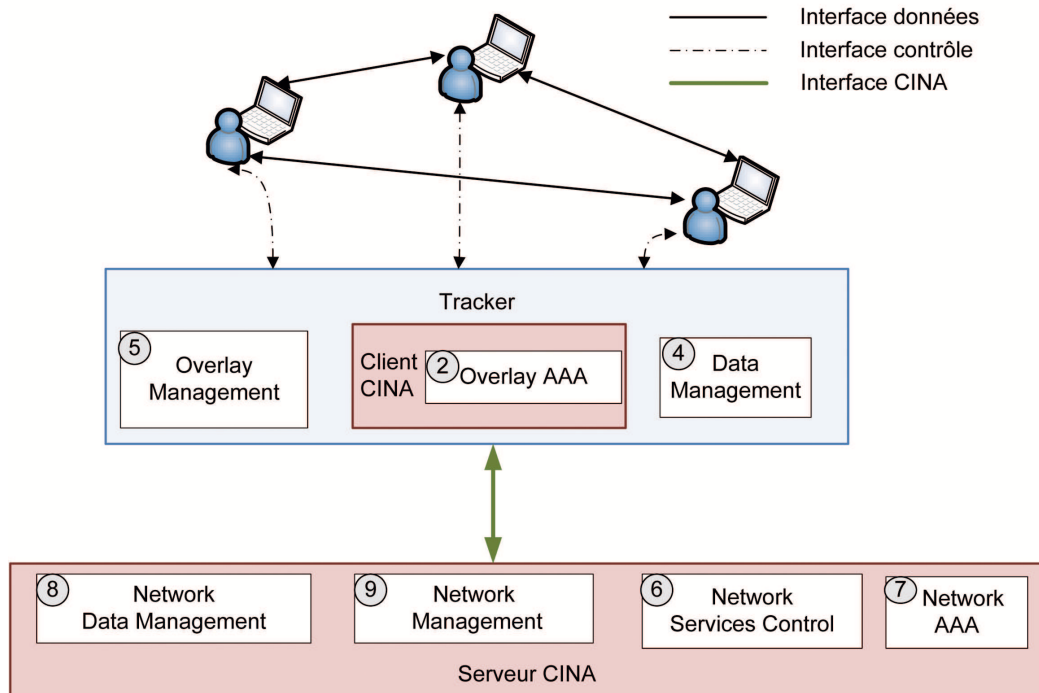


FIGURE 3.3 – Illustration de l'intégration de l'architecture CINA dans un tracker P2P

au niveau du client P2P limite le champ d'intervention à ce client et exclut le reste du réseau *"overlay"*. Par conséquent, il n'est plus possible de fournir des contraintes de service aux opérateurs réseaux car les clients sont totalement indépendants dans leur fonctionnement les uns des autres. D'un autre côté, même une entité centrale de gestion comme un tracker est incapable de synthétiser les besoins et contraintes du réseau P2P. En effet, les échanges de données entre les pairs peuvent être régis par des politiques complexes comme le *"tit-for-tat"* de l'application Bittorent [Coh03]. En conséquence, il est très compliqué de définir une loi caractérisant les échanges de données entre les pairs.

3.2.2 Proposition d'une architecture P2P dédiée

Dans cette section, nous présenterons la nouvelle approche P2P que nous avons conçue. La structuration du réseau P2P permet d'avoir plus de maîtrise sur le trafic de l'*overlay* P2P d'où une meilleure gestion des contraintes et besoins de l'application P2P. Dans cette conception, les pairs sont rassemblés en groupes ou *"clusters"*. Les données sont échangées entre les pairs appartenant au même groupe ou bien à des groupes différents mais suivant des tables de routage prédéfinies.

3.2.2.1 Intérêt de cette conception

La technique que nous proposons présente des avantages indéniables :

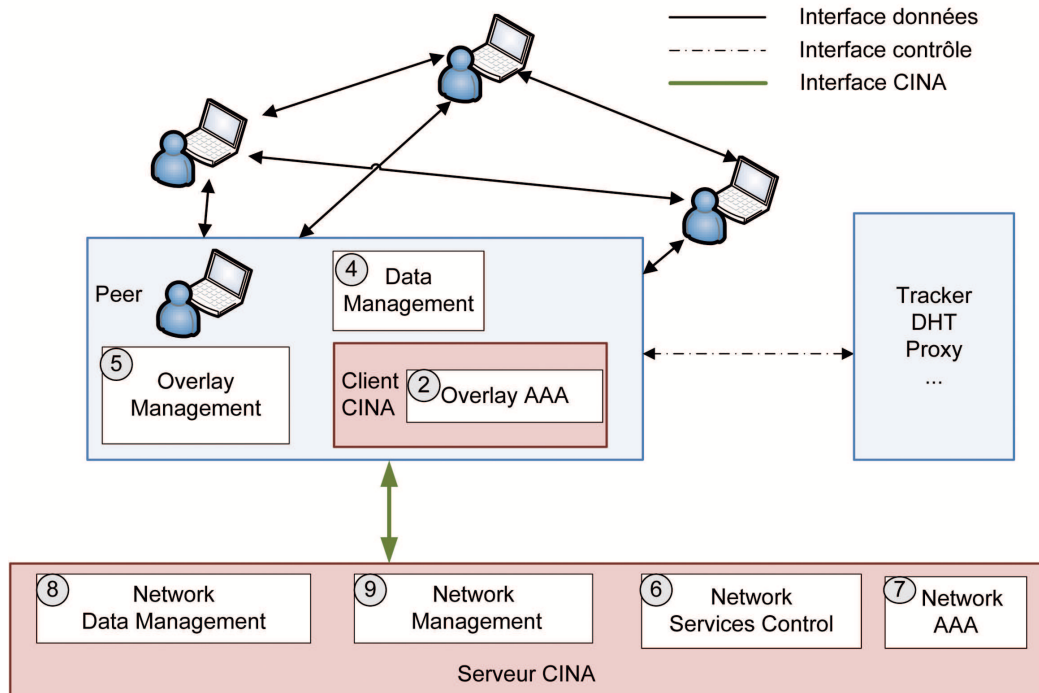


FIGURE 3.4 – Illustration de l'intégration de l'architecture CINA dans un client P2P

- La charge de trafic de l'application P2P est réduite car les messages de contrôle sont contenus. En effet, la conception du système dispense les pairs de s'échanger entre eux les listes de morceaux ou "chunks" dont ils disposent.
- Le délai d'accès aux données est raccourci apportant en conséquence une meilleure efficacité au système. En effet, les chunks sont échangés entre les pairs sans avoir recours au mécanisme de requête-réponse.
- La création et l'interconnexion des clusters P2P sont basées sur les recommandations fournies par les opérateurs de réseaux, ce qui assure une meilleure QoE pour les utilisateurs des applications P2P qui adopteront notre conception.

3.2.2.2 Conception du système

Nous présentons dans cette section la conception du système P2P. Considérons un réseau "overlay" P2P composé d'une ou plusieurs sources S et une population de N pairs. Nous cherchons à définir un plan de routage applicatif **optimal** pour les données provenant des sources S et échangées entre les pairs N . Nous définissons deux structures pour la conception du système. La première concerne la structuration propre aux données. La seconde définit la structure de l'overlay P2P. Notre objectif est de concevoir un système P2P avec des échanges de données structurés, fiables, optimaux et sans recours aux requête-réponse. Nous exposons le système que nous avons défini dans les sections suivantes.

3.2.2.3 Structure des données

La structuration de données que nous proposons les conditionnent en fragments ou "*chunks*" de taille égale. Chaque fragment est identifié par son numéro. Les échanges de fragments entre les pairs sont quant à eux régis par des numéros définis par un schéma d'identification cyclique. Cette identification permet d'attribuer des identifiants à chaque fragment suivant un cycle prédéfini. A titre d'exemple, si on considère que les fragments sont numérotés dans l'ordre et si nous appliquons un cycle de 8, les pairs peuvent déterminer la position de chaque fragment 1 à 8 dans le cycle. Dans ce cas, le fragment dont le numéro est 15 est attribué l'identifiant 7 dans le cycle ($15 \equiv 7 \pmod{8}$) tandis que le fragment qui arrive à la position 33 ($33 \equiv 1 \pmod{8}$) est attribué l'identifiant 1. Cette structure cyclique permet d'organiser les échanges de fragments entre pairs dans les applications P2P et plus particulièrement celles où le nombre de fragments est initialement inconnu comme les applications de diffusion vidéo en direct ou "*live streaming*".

Cette structuration permet d'un côté d'assurer une homogénéité dans le traitement des données pour les clients du réseau P2P et d'un autre côté d'établir une base commune pour un échange piloté des données entre eux. Un échange piloté se traduit par une politique d'échange de données entre les pairs qui soit contrôlée par le système. L'objectif de cette politique est d'atteindre un rendement optimal.

3.2.2.4 Structure de l'overlay

Le second schéma définit la structure de l'*overlay* ou du réseau P2P. Nous proposons d'organiser la population de pairs en groupes de nombre fixe. Chaque cluster reçoit les chunks provenant de l'extérieur (à partir de la source S ou bien de pairs dans d'autres groupes). Les pairs de chaque groupe échangent les chunks entre eux et peuvent en transmettre à d'autres groupes en fonction de la capacité de transmission agrégée du groupe. Dans ce qui suit, nous allons d'abord discuter de la taille des groupes puis nous exposerons l'acheminement des fragments de données entre les différents pairs.

La taille des groupes Dans les travaux de Munding et Al. [MWW08], les auteurs ont démontré que dans un réseau P2P structuré suivant un modèle de transmission/réception simultanée, le nombre minimal de tours requis pour aboutir à une diffusion complète de tous les fragments de données est :

$$M + \lceil \log_2 N \rceil \tag{3.1}$$

M est le nombre de fragments de données et N est le nombre de pair dans l'overlay. Dans leur étude, ils ont assumé que tous les clients disposent de capacités de téléchargement identiques. En plus, ils ont supposé que la source et tous les clients sont connectés entre eux. Nous avons utilisé ce résultat comme une jauge de performance pour le système P2P que nous avons conçu tout en étant conscient que nous ne pourrions atteindre de telles performances à cause des conditions favorables des études dans [MWW08].

Nous ciblons particulièrement l'existence de connections entre tous les clients. En effet, pour une réalisation concrète avec une population de plusieurs milliers de clients, il est impossible d'établir des connexions entre eux. En conséquence, dans notre système, nous nous contentons d'un nombre limité de connexions pour chaque client. La dérivée 3.2 de la loi 3.1 montre que le nombre de tours nécessaires à la diffusion complète des données diminue avec un nombre de connexion croissant.

$$\frac{d(M + \lceil \text{Log}_2 N \rceil)}{dN} = \frac{1}{N \times \ln 2} \quad (3.2)$$

Ainsi pour réaliser un compromis entre un système P2P théorique et optimal et un système réalisable et performant, nous proposons de limiter le nombre de connexions pour chaque client de base à 8. Ce chiffre, qui représente aussi la taille des groupes pour notre système, présente différents avantages :

- Le nombre de connexions maintenues par chaque client est acceptable au regard des machines réelles qui seront clientes de l'application P2P, comme les tablettes et smartphones,
- Le schéma de routage des fragments de données est bien moins compliqué dans un groupe de 8 clients que dans un réseau où tous les clients sont connectés entre eux,
- La formation d'un groupe complet de 8 client est bien plus rapide que celle d'un groupe de 24 clients par exemple. Un groupe complet assure un échange stable et efficace des données

3.2.2.5 L'approche retenue pour les échanges de données

Nous allons présenter dans cette section l'approche que nous avons adoptée pour optimiser le rendement du système. Nous avons défini un échange de données structuré entre les clients sur la base d'une période unitaire de temps T . Pendant, chaque intervalle de temps T , nous allons déterminer l'action réalisée par chaque entité du réseau P2P. Cet intervalle de temps peut représenter par exemple le temps de lecture d'un fragment de données pour une diffusion en direct d'un contenu sur une applications P2P. Dans ce cas, la source devra transmettre une plusieurs copies de chaque fragment de données pendant T aux clients P2P en fonction de sa capacité de téléchargement.

Notre objectif est d'optimiser l'efficacité du système en optimisant l'action de chaque client pendant l'intervalle T . Afin d'atteindre cet objectif, nous avons travaillé sur l'optimisation de deux aspects. Le premier aspect concerne le nombre de fragments échangés entre les pairs pendant l'intervalle de temps T . Plus ce nombre est grand, meilleur est le partage de données dans le réseau P2P et en conséquence son rendement. Le second aspect concerne l'optimisation de l'ordre de transmission des fragments suivant plusieurs critères de priorité. Nous avons retenu particulièrement le retard de transmission des fragments afin d'augmenter leur priorité dans la file d'attente et assurer une meilleure de chance de réception à temps. En conséquence, nous considérons l'efficacité du système comme une composition de deux éléments ; le premier représente l'efficacité en terme de

débit atteint par rapport à la bande passante totale disponible au système par intervalle de temps. Le deuxième élément représente la capacité du système à éviter les pertes de fragments et à les délivrer le plus rapidement possible pour réduire les décalages dans le temps.

3.2.3 L'optimisation du rendement du système

Nous présenterons dans cette section les notations et conditions puis le modèle que nous utilisons pour calculer le rendement du système.

3.2.3.1 Notations et conditions

Nous considérons que chaque client dispose d'un débit suffisant pour la transmission d'au moins un fragment de données à un autre client au cours de l'intervalle T . Cette condition est légitime car le fonctionnement des applications P2P est principalement assuré par la contribution des pairs dans la distribution des données. Ainsi, nous considérons qu'une contribution normale est le partage d'autant de données que la quantité reçue. Nous supposons également que les clients disposent d'une capacité de réception au moins égale à leur capacité de transmission. Nous résumons les notations que nous utilisons dans le tableau 3.1.

TABLE 3.1 – Les paramètres du modèle

Notation	Paramètre
N	Nombre de clients P2P
M	Taille du cycle pour les chunks
C_j	Chunk numéro j
P_j	Priorité de C_j
N_x	Identifiant d'un client dans un groupe $x \in [1, 8]$
UP_x	Capacité de transmission de N_x
DP_x	Capacité de réception de N_x
$R_{x,j}$	Disponibilité du fragment C_j chez le noeud N_x
$T_{x,y,j}$	Transmission du fragment C_j par le noeud N_x au noeud N_y

3.2.3.2 Le modèle d'échange de données

Dans le modèle P2P que nous avons conçu, chaque nouveau client qui rejoint l'*overlay* est affecté à un groupe. Par ailleurs, chaque client joue un rôle bien défini dans le groupe auquel il appartient. En effet, il remplit le rôle d'un noeud spécifique du groupe déterminé par un identifiant de noeud, N_x avec $x \in [1, 8]$.

Réception des données par un groupe Pour éviter un point de défaillance central dans le groupe lorsqu'un seul noeud est responsable de la réception des données de l'extérieur et leurs transmission vers les autres noeuds du groupe, nous proposons une

réception assurée par tous les noeuds. Nous définissons une rotation cyclique entre les noeuds où N_1 reçoit le chunk C_j pendant T_i , puis N_2 reçoit C_{j+1} pendant T_{i+1} , etc.

Priorité des chunks Nous définissons aussi une loi pour déterminer la priorité de chaque chunk au moment de décider lequel transmettre. Chaque fragment est affecté une priorité dans l'ensemble ordonné 1, 10, 20, 30, 40, 60, 70 pour un cycle de $8T$. Chaque fragment débute avec la priorité la plus basse $PTY = 1$ en entrant dans un cluster. Après chaque intervalle de temps T , il récupère la valeur supérieure (voir figure 3.5).

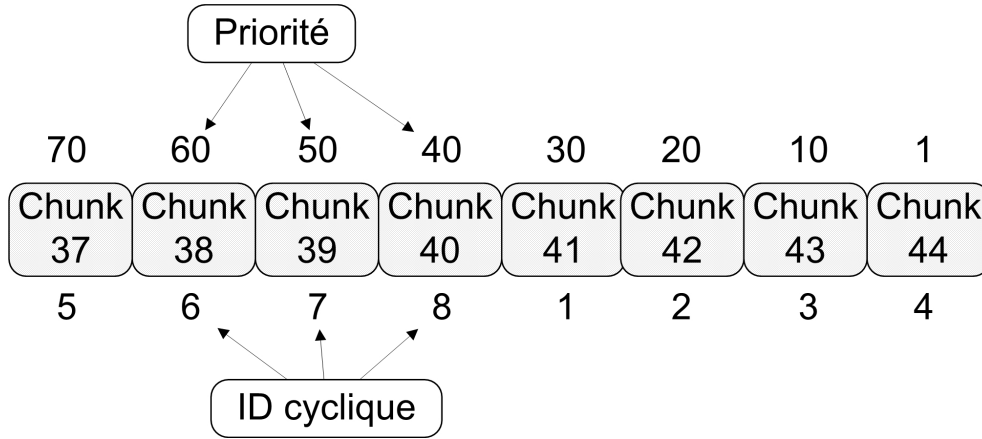


FIGURE 3.5 – Illustration du schéma d'affectation des priorités aux fragments de données

3.2.3.3 L'optimisation du rendement du système

Nous présentons dans cette section la formulation du problème d'optimisation que nous chercherons à résoudre. Nous avons introduit dans la table 3.1 le paramètre $T(x, y, j)$ qui spécifie la transmission du chunk C_j par le noeud N_x au noeud N_y . Ce paramètre prend une valeur binaire égale à 1 lorsque la transmission a eu lieu et 0 sinon. $R(x, j)$ représente la disponibilité de C_j chez N_x , UP_x et DP_x respectivement sa capacité de transmission et de réception. A chaque intervalle de temps T_i , nous calculons l'efficacité du système SE comme la somme de tous les produits des chunks transmis par leur priorité 3.3.

$$SE = \sum_{x=1}^N \sum_{y=1}^N \sum_{j=1}^M T(x, y, j) \times P_j \quad (3.3)$$

Nous cherchons à maximiser SE à chaque intervalle T_i en déterminant quel noeud du groupe doit transmettre quel chunk et à quel autre noeud. Ainsi, la résolution du problème conduit à la détermination de la variable $T(x, y, j)$, tout en satisfaisant les contraintes inhérentes aux :

- capacités limitées de transmission et de réception des clients
- disponibilités des fragments chez les différents noeuds à chaque intervalle de temps

Le problème d'optimisation peut ainsi être exprimé comme suit :

$$\begin{aligned}
\text{Maximiser} \quad & SE = \sum_{x=1}^N \sum_{y=1}^N \sum_{j=1}^M T(x, y, j) \times P_j \\
\text{sujet à} \quad & T(x, y, j) \leq R(x, j) \forall x, y, j \\
& \sum_{x=1}^N T(x, y, j) \leq UP_x \forall x, y, j \\
& \sum_{x=1}^N T(x, y, j) \leq DP_y \forall x, y, j \\
& T(x, x, j) = 0 \forall x, j
\end{aligned} \tag{3.4}$$

La première contrainte stipule que les pairs ne peuvent pas pas transmettre un fragment de données dont ils ne disposent pas encore. Les deux contraintes suivantes encadrent les limites de transmission et de réception des clients par leurs capacité pendant chaque intervalle. La dernière contrainte empêche les auto-transmission de fragments dans le modèle mathématique.

Le problème présenté dans 3.4 est un problème de d'optimisation binaire d'entiers soumis à des contraintes d'égalité et d'inégalité. Nous avons utilisé un algorithme de programmation linéaire de type "*branch-and-bound*" pour la recherche d'une solution optimale au problème [Kia10].

3.2.4 Les échanges de données dans un groupe de pairs

Nous présenterons dans cette section le fonctionnement des groupes P2P. Notons d'abord que la transmission des chunks entre les noeuds répond à un schéma de routage établi. Il permet d'éviter d'inutiles échanges de "*buffer map*" pour indiquer les disponibilités des chunks en mémoire et réduire ainsi la bande passante affectée aux messages de signalisation. Par ailleurs, les délais d'accès aux fragments se trouvent aussi contenus par délaisement des procédés d'exposition de disponibilités et de requêtes/réponses. Au final, rajoutons que le schéma de routage est basé sur l'optimisation de l'efficacité du système au niveau de chaque étape ce qui conduit à des performances intrinsèques intéressantes pour les applications P2P.

Formation des groupes Chaque groupe est formé par un nombre constant de noeuds. Il est construit par les différents clients qui rejoignent l'*overlay* et qui lui sont affectés. Lorsque le nombre de clients dans un groupe atteint la taille de groupe pour l'application, il est considéré comme complet. En conséquence, un nouveau groupe est créé et les nouveaux clients lui sont affectés. Nous allons commencer par exposer le schéma de routage pour un groupe complet, puis nous présenterons les liens entre les groupes.

3.2.4.1 Le schéma de routage pour un groupe complet

Le système P2P que nous proposons définit un rôle spécifique pour chaque noeud dans le groupe. Le rôle détermine exactement quelles seront les actions à exécuter par chaque noeud. Quand un client rejoint le groupe, il est attribué un identifiant de noeud qui spécifie son rôle. Le schéma de routage du système définit pour chaque rôle les chunks, déterminés par leurs identifiants cycliques, qu'il sera le premier dans le groupe à recevoir de l'extérieur. Ainsi par exemple, dans un schéma de routage peut particulier, N_1 sera le noeud du groupe à recevoir le chunk C_1 à T_1 . Ensuite, N_2 recevra C_2 à T_2 et ainsi de suite. Nous supposons que dans cet exemple, tous les clients ont la capacité de télécharger au moins un fragment de données par intervalle de temps.

Le calcul des tables de routage Pour définir le schéma de routage, nous avons exécuté l'algorithme présenté dans 3.4. Pendant et après chaque intervalle de temps, nous mettons à jour les capacités de transmission et de réception des noeuds et la liste des chunks dont ils disposent en prenant en compte aussi les fragments qu'ils téléchargent à partir de sources externes au groupe. La résolution du problème d'optimisation sou les contraintes énoncées dans 3.4, détermine le schéma de routage optimal pour les fragments de données entre les clients. Ce schéma maximise l'efficacité du système. Nous avons ainsi pu définir des tables de routage.

Le Régime transitoire et le régime périodique Le schéma de routage pour un groupe de noeuds est régi par deux régimes ; transitoire précédant le périodique. Le régime transitoire accompagne tout changement dans le groupe comme l'arrivée d'un nouveau client par exemple. Il s'en suit un régime périodique où la table de routage est statique. Ainsi l'action de chaque noeud pendant chaque intervalle de temps est bien prédéfinie. Un exemple de régime transitoire est illustré dans la figure 3.6. Les six premiers intervalles de temps après la formation d'un groupe complet sont présentés. Par exemple, au cours de T_1 , le noeud N_1 reçoit C_1 d'un noeud X qui n'appartient pas au groupe. Ensuite, pendant T_2 , N_1 devient le seul noeud qui dispose d'un fragment de données à partager que les autres n'ont pas. L'optimisation des échanges de données établit que C_1 soit transmis à N_3 pendant T_2 . Dans le même temps, N_2 reçoit C_2 à partir d'un noeud externe X .

Le régime stable pour un groupe de 8 noeud et un cycle de taille 8 commence à T_{25} , c'est-à-dire après 3 cycles en régime transitoire. La table de routage qui sera utilisée pendant le régime périodique est donc aussi périodique avec une période $8T$, c'est-à-dire que les noeuds répètent les mêmes actions tous les cycles (illustrée dans la table 3.2). Ainsi, à chaque intervalle de temps T , chaque noeud sait exactement quel chunk envoyer à quel autre noeud. Par exemple dans la table , à T_1 , N_1 doit envoyer C_4 à N_7 .

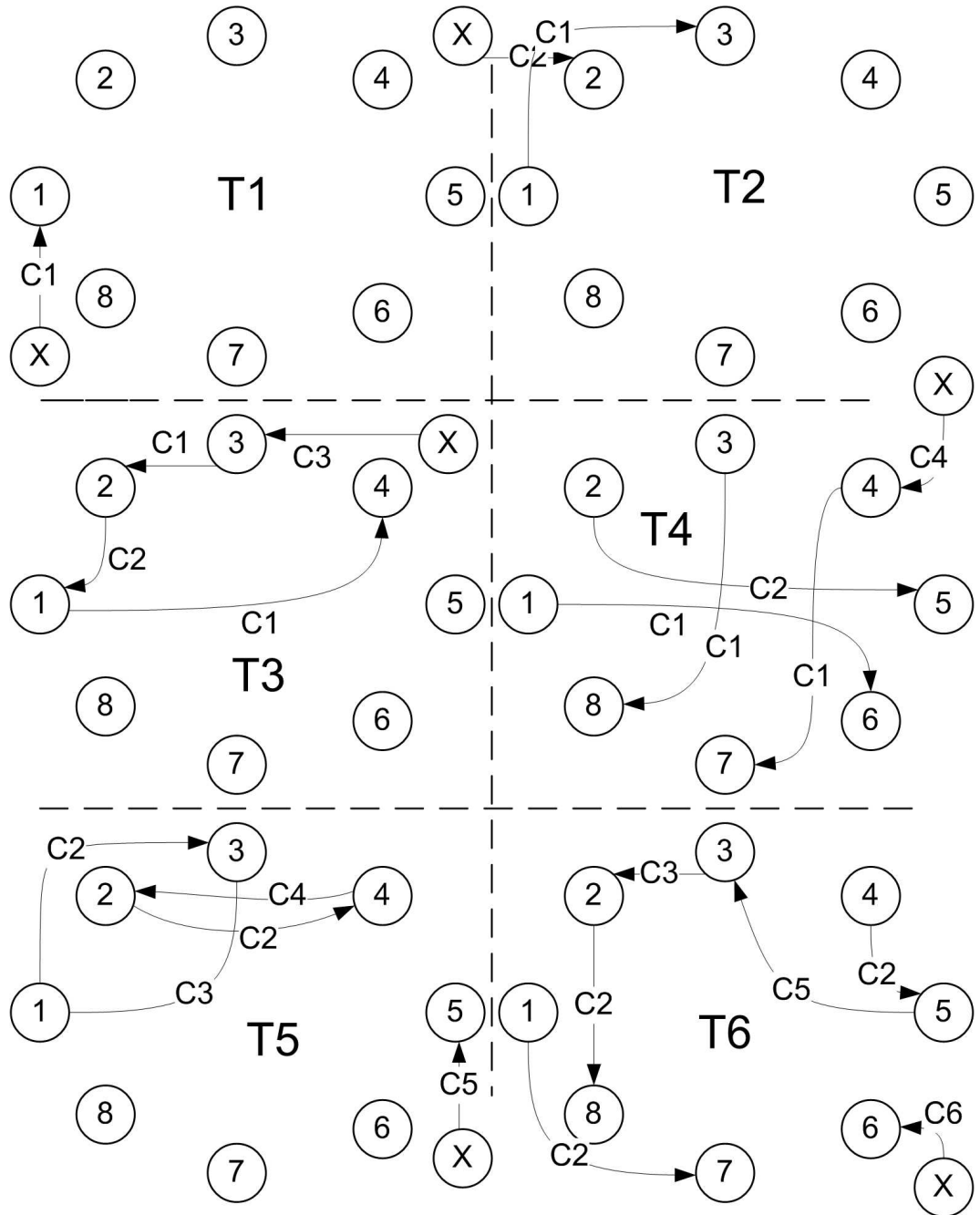


FIGURE 3.6 – Illustration des six premiers intervalles de temps du régime transitoire après la formation d'un groupe complet de huit noeuds.

TABLE 3.2 – La table de routage pour le régime périodique d'un groupe de taille 8 pour un cycle de taille 8

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8
X	$C_1 \rightarrow N_1$	$C_2 \rightarrow N_2$	$C_3 \rightarrow N_3$	$C_4 \rightarrow N_4$	$C_5 \rightarrow N_5$	$C_6 \rightarrow N_6$	$C_7 \rightarrow N_7$	$C_8 \rightarrow N_8$
N_1	$C_4 \rightarrow N_7$	$C_4 \rightarrow N_8$	$C_1 \rightarrow N_4$	$C_1 \rightarrow N_6$	$C_8 \rightarrow N_4$	$C_2 \rightarrow N_4$	$C_3 \rightarrow N_2$	$C_3 \rightarrow N_6$
N_2	$C_5 \rightarrow N_3$	$C_5 \rightarrow N_4$	$C_5 \rightarrow N_7$	$C_2 \rightarrow N_5$	$C_2 \rightarrow N_1$	$C_8 \rightarrow N_3$	$C_4 \rightarrow N_1$	$C_3 \rightarrow N_7$
N_3	$C_4 \rightarrow N_6$	$C_5 \rightarrow N_1$	$C_5 \rightarrow N_8$	$C_6 \rightarrow N_8$	$C_7 \rightarrow N_8$	$C_3 \rightarrow N_1$	$C_3 \rightarrow N_4$	$C_3 \rightarrow N_5$
N_4	$C_4 \rightarrow N_5$	$C_6 \rightarrow N_7$	$C_7 \rightarrow N_1$	$C_1 \rightarrow N_7$	$C_1 \rightarrow N_3$	$C_4 \rightarrow N_2$	$C_2 \rightarrow N_3$	$C_4 \rightarrow N_3$
N_5	$C_6 \rightarrow N_2$	$C_5 \rightarrow N_6$	$C_7 \rightarrow N_2$	$C_8 \rightarrow N_2$	$C_8 \rightarrow N_6$	$C_2 \rightarrow N_7$	$C_2 \rightarrow N_8$	$C_5 \rightarrow N_2$
N_6	$C_6 \rightarrow N_5$	$C_6 \rightarrow N_3$		$C_7 \rightarrow N_3$	$C_1 \rightarrow N_2$	$C_1 \rightarrow N_5$	$C_6 \rightarrow N_5$	$C_6 \rightarrow N_1$
N_7	$C_3 \rightarrow N_8$	$C_7 \rightarrow N_5$	$C_7 \rightarrow N_6$			$C_1 \rightarrow N_8$	$C_2 \rightarrow N_6$	$C_7 \rightarrow N_4$
N_8			$C_8 \rightarrow N_5$	$C_8 \rightarrow N_1$	$C_8 \rightarrow N_7$			

Ce système permet aux applications P2P d'éviter d'utiliser des messages de signalisation entre pairs pour échanger les disponibilités des fragments données et par la suite d'envoyer des demandes de transmission. En effet, chaque noeud du groupe dispose d'informations sur la disponibilité des chunks chez les autres. En outre, les transmissions de fragments de données sont régies par les tables de routage et ne nécessitent pas de requêtes ou d'interventions de la part des clients.

3.2.5 La formation d'un overlay de groupes P2P

Nous pouvons remarquer dans la table 3.2, qu'à chaque intervalle de temps, il y a des ressources de transmission qui sont inexploitées (une case vide pour chaque colonne). En effet, la réception de tous les chunks par tous les noeuds du groupes pendant une période de $8T$ nécessite $64 = 8(\text{receptions}) \times 8(\text{intervalles})$ opérations. Cependant, à chaque intervalle de temps, un noeud X externe au groupe contribue en injectant un nouveau chunk dans le groupe. Ainsi, il suffit que 7 noeuds du groupe contribuent au partage de données pour assurer la transmission de tous les chunks à tout le groupe. En conséquence, des ressources sont donc libérées. Elles peuvent être exploitées pour transmettre des fragments de données à un autre client ou un autre groupe. La formation du réseau P2P global est ainsi établie sur la base de connexions parent-descendant entre les groupes de clients qui se forment.

3.2.5.1 Exploitation des ressources additionnelles

Nous avons déjà évoqué la nécessité de disposer d'une capacité de minimale permettant la réception et la transmission d'un chunk par intervalle de temps pour chaque client. Nous avons justifié cette condition par le mode de fonctionnement intrinsèque des applications P2P. Cependant, certains clients peuvent disposer de capacités plus élevées. Nous n'introduisons pas de notion de super-pairs dans notre modèle car il n'y a rien de fonctionnel qui les différencient des autres clients. Par contre, ces noeuds peuvent contribuer avec plus de ressources dans la transmission des données aux autres. Cette différence peut avoir comme origine par exemple le types d'accès au réseau de chaque client. Ainsi, ceux qui disposent d'un accès à Internet par la fibre optique Fiber To The Home (FTTH) possède des capacités de transmission bien plus supérieure que celle d'une ligne Asymmetric Digital Subscriber Line (ADSL). En fonction de leurs capacités de transmission, les clients peuvent contribuer différemment à l'application P2P. Pour exploiter ces ressources, nous introduisons le paramètre Capacité Unitaire de Transmission (CUT), qui est la capacité de transmettre un fragment de données par intervalle de temps. Certains clients peuvent ne pas disposer d'assez de ressources pour assurer une CUT dans des situations particulières, comme par exemple des clients P2P connectés via un accès mobile. D'autres utilisant un accès FTTH ou ADSL peuvent facilement disposer d'une capacité de 3 ou 4 CUTs à mettre au service de l'application P2P.

Ces CUTs additionnels permettent à chaque groupe de transmettre des chunks à un ou plusieurs autres groupes de l'overlay. Par exemple, si le groupe A est composé de 7 noeuds avec une capacité de 1 CUT chacun et un autre noeud avec une capacité de 2 CUTs, alors ce groupe est en mesure d'envoyer des chunks simultanément à deux autres groupes de l'overlay. Nous déterminons le nombre de groupes descendants qu'un groupe spécifique G_u de taille 8 est capable de servir en tant que parent dans l'équation 3.5.

$$G_u^d = \sum_{x=1}^N UP_x - 7 \quad (3.5)$$

Nous présentons les performances de notre modèle pour certains cas d'application dans la table . Les performances sont comparées aux valeurs théoriques optimales calculées par Mundinger et al. [MWW08] dans leur modèle qui assume la possibilité d'interconnecter tous les pairs de l'overlay. Nous nous sommes arrêtés à un nombre de 4 groupes descendants pour limiter les perturbations dans l'overlay quand il y a des changements dans le grand parent. Les performances obtenues sont intéressantes par rapport aux références théoriques. Nous observons aussi que la durée du régime transitoire est bien contenue. Elle peut atteindre une durée de $48T$ qui consiste en 1 seconde lorsque $T = 50ms$. La période du régime stable est très courte (entre $8T$ et $16T$) ce qui offre l'avantage d'avoir des tables de routage petite de taille.

TABLE 3.3 – Comparaison des performances de notre modèle par rapport aux limites théoriques d’optimalités. Les groupes P2P ont une taille de 8 noeuds avec des capacités de transmission différentes et un nombre de groupes descendants différents

Capacité de transmission agrégée du groupe	Nombre de groupes descendants	Durée du régime transitoire	Durée du régime périodique	Durée d’échanges du modèle (T)	Durée de référence théorique (T)
8	0	24T	8T	13	11.16
	1	48T	8T	14	
9	0	16T	8T	12	11.32
	1	16T	8T	12	
	2	32T	8T	13	
10	3	32T	16T	13	11.45
11	4	32T	16T	13	11.58

3.2.5.2 Atténuation des instabilités de groupes dans l'overlay

Les applications P2P font face en pratique à un phénomène d'instabilité engendrée par une classe de clients qui rejoignent et quittent de façon brève l'overlay. Ce phénomène peut induire des problèmes d'instabilité dans les groupes P2P dûs aux changements fréquents du nombre de noeuds dans chaque groupe. Afin d'atténuer ce phénomène, nous introduisons un noeud supplémentaire particulier à chaque groupe, que nous appelons le Pair de Stabilité (PS). Nous lui attribuons l'identifiant de noeud N_9 dans le cluster.

N_9 reçoit les chunks en priorité par rapport aux autres noeuds du cluster. Par exemple, si N_1 reçoit le chunk C_1 à T_i d'une source X externe, à T_{i+1} , N_1 transmettra C_1 à N_9 (voir figure 3.7). Par ailleurs, N_9 n'est pas sollicité pour contribuer à la distribution des données pour le groupe. Cette approche garantit ainsi que N_9 soit capable d'assumer le rôle de n'importe quel autre noeud du groupe car il reçoit les fragments de données avant eux. En conséquence, N_9 pourra substituer instantanément et de façon fiable n'importe quel client de son groupe qui quitterait l'application. Le groupe est ainsi en mesure de garder le même schéma de routage et bénéficier d'une meilleure stabilité. En parallèle, un nouveau pair est affecté au groupe pour prendre la relève de ce rôle de N_9 .

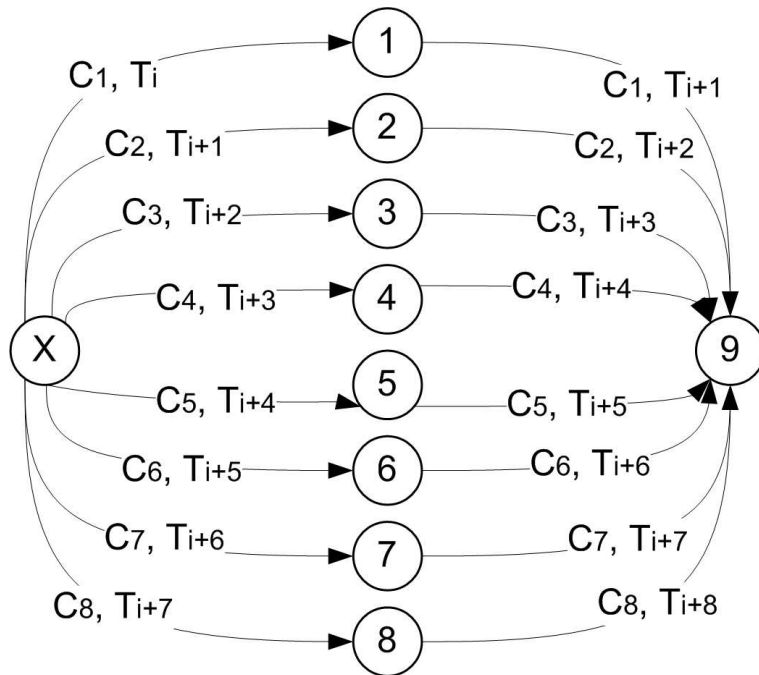


FIGURE 3.7 – Illustration des échanges de données en présence du pair de stabilité. Les fragments de données sont transmis en priorité à ce noeud particulier, identifié N_9 dans chaque groupe.

Cette technique permet de réduire le phénomène d'instabilité quand son ampleur est limitée, c'est-à-dire quand les arrivées et départs des clients sont raisonnables. Dans des conditions sévères, les groupes P2P devront permuter vers des schémas de routage propres à des groupes incomplets constitués de 7 noeuds voir moins jusqu'à ce que de nouveaux membres leurs soient affectées. Cependant, nous croyons que des conditions sévères d'arrivées et départs fréquents de clients affecteraient toute application P2P même avec un fonctionnement différent. Notre système dispose de cet argument de pouvoir maîtriser le phénomène dans des conditions raisonnables afin d'assurer une bonne QoE pour les utilisateurs.

Exploitation des ressources de N_9 Les capacités de transmission dont dispose le noeud N_9 dans chaque groupe peuvent être utilisées pour envoyer des fragments de données à d'autres groupes, et ceci est d'autant plus intéressant que N_9 reçoit les données immédiatement après leur réception dans son groupe. En conséquence, les destinataires des données transmises par N_9 les recevront encore plus vite. Lorsque N_9 est appelé à remplacer un noeud qui a quitté le groupe, les actions de transmission qui lui étaient destinées seront utilisées pour servir les noeuds auxquels ils transmettait des chunks, ainsi la continuité de service est préservée jusqu'à ce qu'un nouveau client rejoigne le groupe pour reprendre ce rôle.

Coût de l'ajout du noeud N_9 Le mécanisme que nous proposons est accompagné d'un coût en terme de performance que nous jugeons acceptable (voir table 3.4). En effet, le fait d'imposer la transmission prioritaire de tous chunks à N_9 avant les autres noeuds et interdire simultanément à N_9 de contribuer aux échanges dans son groupe rallonge la durée moyenne d'accès aux données. Au final, nous observons que l'acheminement de tous les fragments de données à tous les noeuds d'un groupe nécessite un intervalle de temps supplémentaire. Ce retard est compensé en contre-partie par une meilleure résistance de l'application aux phénomènes d'instabilité.

TABLE 3.4 – Comparaison des performances du modèle avec ou sans l'utilisation d'un pair de stabilité

	Groupe sans PS	Groupe avec PS
Nombre moyen d'intervalles T pour la réception d'un chunk par tous les noeuds	3.62	4.66
Nombre maximal d'intervalles T pour la réception de n'importe quel chunk par tous les noeuds	5.75	7.43
Durée du régime transitoire	$32T$	$32T$
Périodicité du régime stable	$8T$	$16T$
Durée pour la réception de tous les chunks par les noeuds	$13T$	$14T$

3.2.6 La construction de l'overlay P2P

Les performances des applications P2P dépendent aussi directement de la construction de leur topologie. Nous avons évoqué dans les sections précédentes l'utilité des interfaces comme ALTO et CINA dans la construction de la topologie de l'overlay. Le choix de la liste des pairs auxquels sera connecté chaque client détermine la qualité de réception de données pour chaque client. Dans le modèle que nous proposons, les performances de transport dépendent des connexions intra et inter groupes, c'est-à-dire du choix des pairs pour former un groupe puis des choix de groupes parents et descendants. Nous présentons dans cette section les approches que nous adoptons pour la construction d'un maillage efficace et performant pour l'overlay P2P.

3.2.6.1 Connexions intra-groupe

L'affectation des clients dans des clusters est régie un critère de proximité géographique ou réseau. Cette proximité permet d'atteindre de meilleures performances grâce à la réduction de la distance qui s'accompagne généralement d'une réduction de la latence. Ce genre d'informations peut être fourni par les opérateurs réseaux à travers des cartes réseaux. Ainsi, les clients qui sont dans le même PID, sont regroupés ensemble en clusters. Quand il y a besoin de requérir à des clients dans des PIDs différents pour former un cluster, d'autres critères comme la latence, la distance ou bien le coût de transport peuvent être employés.

3.2.6.2 Connexions entre groupes dans le même PID

Les données sont échangées entre les clusters à travers des connexions établies entre clusters parents (source) et descendants (destination). Nous présentons dans cette section les approches pour connecter des groupes qui sont dans le même PID.

Les pairs qui sont dans le même PID sont regroupés ensemble dans des clusters. En conséquence, chaque PID pourrait contenir un certain nombre de clusters. Nous appliquons la même approche et privilégions les connexions entre ces groupes qu'avec des groupes dans d'autres PIDs. Cependant, deux aspects sont à prendre en considération pour garantir un système fiable.

Le premier aspect concerne l'isolement des groupes dans un PID. Les connexions intra-PID peuvent aboutir sur des îlots totalement isolés du reste de l'overlay. Dans ce cas, les groupes ne sont pas en mesure de récupérer les données de l'extérieur. La solution que nous proposons définit une arborescence de groupes à l'intérieur de chaque PID avec un groupe à la racine qui assure la réception des données de l'extérieur de son PID et qui les transmet aux autres groupes.

Le deuxième aspect concerne la performance des arborescences construites dans chaque PID. En effet, un nombre important de groupes dans un arbre entraîne un délai supplémentaire pour la réception des données qui sont transmises des groupes parents aux

groupes descendants. Par ailleurs, une défaillance d'un groupe de l'arbre affecte immédiatement ses descendants. En contre-partie, un nombre réduit de groupes résulte en charge de trafic plus importante. En conséquence, nous avons établi des règles pour la construction des arbres dans notre système en fixant deux limites : une profondeur maximale et un nombre maximal de groupes dans chaque arborescence. Ces limites traduisent un compromis entre les performances de l'application et sa charge de trafic sur le réseau.

Nous exposons dans la figure 3.8 deux constructions différentes d'arborescence. L'arbre à gauche contient 20 groupes avec une profondeur de 4 niveaux. Ceux qui sont à droite sur la figure contiennent autant de groupes mais limitent le nombre maximal à 7 pour une profondeur maximale de 3 niveaux. Alors que dans la première construction, un seul cluster est responsable de la réception des données pour l'ensemble des autres clusters, dans la seconde construction, les 3 clusters parents engendrent plus de trafic inter-PID. En contre-partie, ils offrent de meilleures performances grâce à un acheminement des données plus rapide et plus fiable.

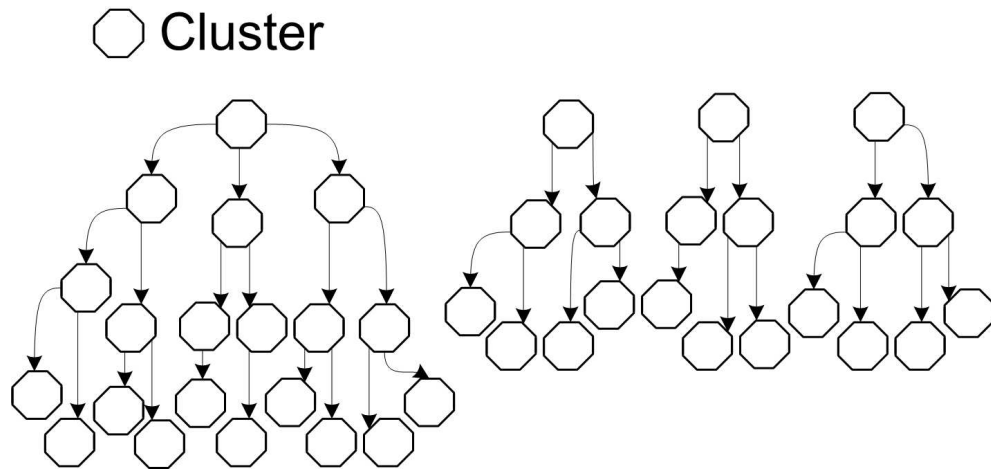


FIGURE 3.8 – Illustration de deux constructions d'arborescence avec des tailles et des profondeurs d'arbre différentes.

3.2.6.3 Connexions entre groupes dans des PIDs différents

Les arbres de groupes construits par le système que nous proposons nécessitent la réception de données à partir de sources externes à leur PID. Cette condition évite la formation d'îlots de groupes isolés du reste de l'overlay. Afin de choisir le meilleur parent parmi la liste de source disponible pour servir chaque arbre, nous exploitons des informations qui peuvent être fournies par l'interface CINA sur la distance, la latence ou bien le coût d'acheminement pour déterminer le meilleur chemin réseau entre la racine de l'arbre et la liste de sources retenue.

Chapitre 4

Evaluations des performances des approches de collaboration

Afin d'étudier les performances de l'interface de collaboration CINA, nous nous sommes appuyés sur plusieurs solutions. Nous avons développé des modèles de simulation analytiques pour les applications d'une importante étendue. Nous avons aussi réalisé des expérimentations réelles sur des prototypes concrets. Nous avons contribué par ailleurs à la validation du fonctionnement de l'interface conjointement dans le projet ENVISION [BM11a, BM11b]. Notre principal objectif cependant est d'évaluer l'utilité de l'interface CINA à travers certains cas d'application et observer les performances obtenues sous différentes conditions. Dans ce chapitre, nous présentons l'implémentation de notre modèle de simulation ainsi que les différentes évaluations que nous avons menées pour des services particuliers. Ensuite, nous exposerons les détails du prototype que nous avons développé et les tests que nous avons effectués avant de conclure.

4.1 Les simulations analytiques

Nous avons développé un modèle de simulation pour les réseaux à commutation de paquets à partir de travaux existants sur la théorie des files d'attente. Nous exposerons dans cette section les raisons qui nous amènent à utiliser ce modèle avant de le présenter en détail.

4.1.1 Les principaux défis relatifs aux évaluations

L'exploitation des interfaces de collaboration de type CINA et ALTO nécessite l'accès à des informations relatives à la fois aux réseaux et services. Côté réseaux par exemple, il est impératif de disposer de données sur la topologie du réseau, ses caractéristiques, les plans et les politiques de routage, etc. En effet, ces informations sont nécessaires par exemple à la construction de la carte réseau et de la matrice des coûts. En conséquence, toute évaluation d'un algorithme qui exploite des informations réseaux est voué à l'échec sans la maîtrise de l'outil réseau. De manière similaire, les traces et

l'historique de trafic réseau sont tout simplement inutilisables si on n'a pas accès aux informations sur le réseau de transport.

complexité des expérimentations à large échelle La meilleure façon d'évaluer les performances des algorithmes reste incontestablement l'expérimentation dans des réseaux physiques. Une possibilité serait de mener des tests dans des réseaux à étendue locale. Cependant, elle ne permet pas d'aboutir à des résultats concluants. En effet, outre le fait que la diversité de chemins et de serveurs est très limitée, le nombre de clients admissibles n'est généralement pas représentatif. Sans citer le fait que les résultats peuvent être facilement obtenus analytiquement. Une autre possibilité serait de réaliser des expérimentations dans un réseau opérationnel. Cette gamme assure des tests à large échelle dans des conditions réelles. Cependant, cette tâche présente d'énormes défis très compliqués. En effet, elle nécessite l'accès aux infrastructures d'un réseau opérationnel, la manipulation d'importantes quantités de données, la possibilité de déployer ou bien d'utiliser des services existants, CDN ou P2P, la participation d'une population importante de clients. Réunir toutes ces conditions semble compromis d'avance. A notre connaissance, il existe aujourd'hui deux expérimentations de cette envergure [GLP⁺09, LG12]. Chaque expérimentation était portée par une coopération de grands acteurs (Comcast et Pando pour la première et China Telecom et Xunlei pour la seconde). Par ailleurs, ces tests ont été menés pour évaluer les performances d'une application P2P conjuguée à l'utilisation de l'interface ALTO. Ils se sont appuyés sur des coûts prédéfinis pour remplir la matrice de coûts. Ceci confirme d'autant plus la difficulté de mettre en place des tests avec des mesures dynamiques dans un réseau réel. Il s'agit bien sûr d'abord d'avoir accès aux infrastructures puis de développer les modules nécessaires.

Ainsi, une approche plus raisonnable pour évaluer les performances de nos propositions est donc de s'appuyer sur des simulations. En effet, lorsque les expériences réelles s'avèrent compliquées, il est plus sage de valider les propositions par des résultats analytiques avant d'entamer des tests réels. Dans la section suivante, nous présentons le modèle de notre simulation.

4.1.2 Introduction au modèle

Dans les modèles intégrant des centaines voire des milliers de nœuds, les simulations continues dans le temps sont très exigeantes car elles nécessitent énormément de ressources de calcul et de temps [Liu02]. Pour parvenir à outrepasser ces difficultés, nous proposons l'utilisation d'un modèle de simulation analytique qui s'appuie sur les réseaux de files d'attente multi-classes de type Baskett, Chandy, Muntz, Palacios (BCMP) [BCMP75] pour modéliser le réseau et le trafic des applications. Le modèle nous permet d'évaluer les performances de divers paramètres relatifs au réseau et au trafic. Le modèle BCMP est un réseau de files d'attente convenable pour la modélisation de réseaux étendus à commutation de paquets [Nai98, Kle76, GK77, LW81]. Il a aussi été utilisé pour la modélisation de divers autres systèmes comme pour la surveillance du

trafic routier [DOL11], l'étude de l'équité entre les classes [LW81] ou l'étude de l'unicité d'un routage optimal entre une source et une destination dans le contexte de la disponibilité de plusieurs chemins avec des classes différentes [AK01]. Nous nous intéressons à la faculté de ce modèle d'analyser indépendamment différentes classes dans un réseau à grande échelle ce qui nous permettra d'évaluer les performances de différents services et applications simultanément. Le modèle est basé sur un système discret dont les paramètres et les performances peuvent respectivement être modifiés et mesurés de façon discrète dans le temps.

4.1.2.1 Notations

Nous présentons dans cette section les différentes notations et assumptions retenues pour notre système.

Le réseau Le modèle de réseau de files d'attente que nous employons un modèle de réseau ouvert, c'est à dire avec des arrivées et des départs, sur la base d'un graphe orienté de N sommets et M arcs. Chaque sommet modélise un PID du réseau. Rappelons qu'un PID peut représenter une adresse IP, un sous-réseau, un ensemble de sous-réseaux voir même un système autonome ou une entité supérieure. Les arcs modélisent quant à eux les liens du réseau. Leurs poids représentent la capacité de chaque lien. Il est possible aussi dans le modèle de définir des distances pour prendre en compte le temps de propagation sur des liens de longue distance. Le routage dans le réseau est statique et est indépendant de l'état des liens. Il est calculé à travers l'algorithme de Dijkstra à partir des poids associés aux différents liens (c'est-à-dire leurs capacités). Nous supposons que le délai de traitement des paquets sont constants et négligeables car généralement insignifiants comparés aux temps de transmission et propagation dans les réseaux IP. Sur la base de ces considérations, on peut maintenant modéliser un réseau de commutation de paquets à l'aide d'un réseau de files d'attente, où les liens sont représentés par des stations de file d'attente et les paquets par les clients du réseau (voir figure 4.1).

Le lien est représenté par une file d'attente de transmission et une file d'attente de propagation. Cette dernière est de type serveur infini, c'est-à-dire qu'il n'y a pas de temps d'attente dans la file ce qui caractérise la propagation. Les paquets, qui sont les clients du réseau, sont traités au niveau de chaque sommet ou station pour être orientés vers un autre sommet ou bien quitter le réseau.

Le trafic dans le réseau La modélisation du trafic dans le réseau est assujettie à un certain nombre de lois. Le processus d'arrivée des paquets est un processus de Poisson d'intensité λ_r pour chaque classe r de trafic. La taille des paquets est définie par une distribution exponentielle avec une moyenne $\frac{1}{\mu}$, dont la fonction de densité est $\mu \exp(-\mu x)$ pour $x \geq 0$. Dans les simulations que nous avons exécutées, une partie du trafic provient des services que nous étudions dans le cadre de la collaboration avec les réseaux de transport. L'autre partie provient du reste des services et applications qui exploitent le réseau simultanément. Les paramètres de ce trafic sont aléatoirement

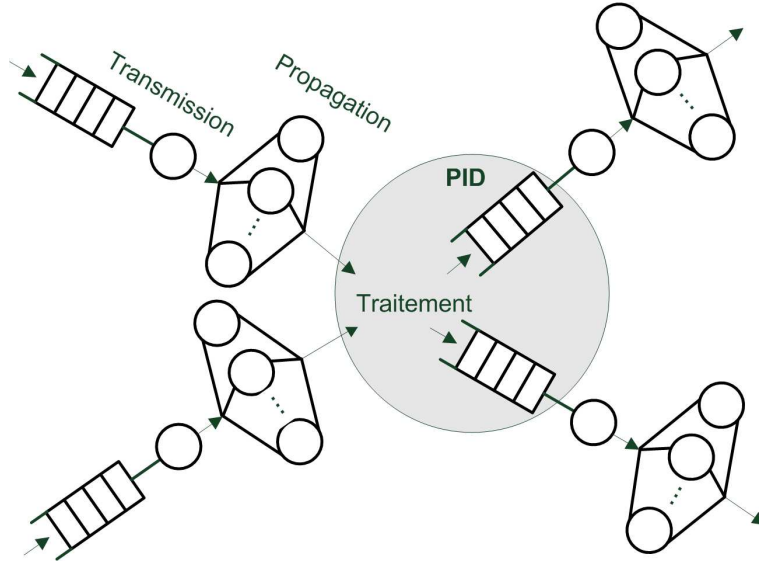


FIGURE 4.1 – Illustration de la modélisation d’un noeud d’un réseau à commutation de paquet par les files d’attentes

choisi. Ainsi les intensités de trafic λ_r sont générées de manière aléatoire dans des plages pré-établies. Aussi, les probabilités d’arrivée des paquets à chaque station et les probabilités de transition d’une station à une autre sont également générées de manière aléatoire pour chaque classe de trafic. Ces probabilités sont définies dans une matrice de probabilité de transition $[M, M]$.

Les applications Les différentes applications étudiées sont régies par un certain nombre de paramètres que nous faisons évoluer aléatoirement dans des plages pré-établies durant nos simulations. Ainsi le nombre de clients ou bien le nombre d’instance de service sont aléatoirement définis. De la même manière, la répartition de ces entités entre les différents PID du réseau évolue de manière aléatoire.

La génération aléatoire de paramètres Certains paramètres de configuration sont difficiles à avoir car portant des informations confidentielles ou jugées privées, d’autres sont compliqués à manipuler alors que l’utilité d’autres est dépendante d’autres paramètres. Pour toutes ces raisons et celles évoquées précédemment, nous avons axé notre modèle sur l’évolution aléatoire des valeurs des paramètres d’une simulation à une autre avec une exécution suffisamment. Les résultats ne traduiraient pas une situation particulière mais plutôt une représentation efficace d’un ensemble d’exécutions.

4.1.3 Présentation du modèle

Nous présentons le modèle analytique et sa configuration pour les simulations que nous avons effectuées dans les paragraphes qui suivent.

4.1.3.1 Les conditions du réseau BCMP

Le modèle repose sur un réseau de files d'attente de type BCMP. Ce type de réseau offre la possibilité de réaliser des mesures de performances pour des réseaux ouverts (qui est la catégorie de réseau nous intéressent), fermés ou mixtes. Certaines conditions sont à respecter pour les réaliser :

- Les taux d'arrivées de paquets dans le réseau doivent être régis par une distribution de Poisson de paramètre λ_r avec $r \in [1, R]$ et R le nombre de classes dans le réseau
- Les taux de services μ_i , $i \in [1, M]$ où M est le nombre de stations (liens réseaux) doivent suivre une loi exponentielle.
- La discipline de service pour le traitement des paquets doit être conforme à l'une des politiques suivantes : premier arrivé premier servi, partage des ressources, service immédiat ou bien dernier arrivé premier servi.

Ces conditions maintiennent les propriétés nécessaires pour aboutir à des processus indépendants et identiquement distribués. Cette approche permet la résolution des états d'équilibre du réseau par les outils employés pour les processus de Markov à temps continu [LW81]. Nous avons retenu le modèle de réseau ouvert BCMP pour nos simulations pour permettre l'arrivée de paquets depuis des sources externes comme les noeuds de distribution CDN, instance data-center ou bien source P2P. La discipline de service est "premier arrivé, premier servi" car nous ne nous intéressons pas aux politiques de différenciation de traitement des paquets. Nous supposons par ailleurs que les équipements du réseau sont capables de stocker tous les paquets en voie de traitement, ce qui assuré dans des conditions normales. Enfin, la dernière condition stipule que les changements de classe ne sont pas autorisés ce qui respecté car chaque classe représente un ou plusieurs applications avec un trafic indépendant distingué.

4.1.3.2 Les paramètres du modèle analytique

Les paramètres du modèles sont résumés dans la table 4.1.

4.1.3.3 La stabilité du système

La vérification de la stabilité du système est une phase nécessaire avant l'entame des calculs relatifs à ses états et à ses performances. commençons par les probabilités introduites dans la table 4.1 et qui satisfont les relations suivantes :

$$\sum_{i=1}^M p_{ij,r} = 1 \forall i \in [0, M] \quad (4.1)$$

$$\lambda_{ir} = e_{ir} \times \lambda_r \quad (4.2)$$

L'équation 4.1 précise que la somme des probabilités d'arrivées des paquets à la station j depuis toutes les autres stations et depuis l'extérieur du réseau est égale à 1. L'équation 4.2 précise que le taux d'arrivée des paquets d'un flux de trafic r dans une station i est égale aux taux de visite du flux à la station multiplié par son intensité.

De ces deux équations, nous pouvons considérer que le modèle du réseau présente les mêmes caractéristiques qu'une file d'attente de type $M/M/1$ avec un taux d'arrivée de loi Poisson de paramètre λ_{ir} , de taux de service de loi exponentielle de paramètre μ_i et de discipline "premier arrivé, premier servi". Les mêmes conditions de stabilité s'appliquent donc sur chaque file d'attente du réseau, à savoir le trafic λ_i sur chaque lien i du réseau doit être inférieur à sa capacité μ_i :

$$\lambda_i < \mu_i \forall i \in [1, M] \quad (4.3)$$

La stabilité du réseau est assurée par la stabilité de l'ensemble de ses stations. Les taux d'arrivée λ_i sont calculés à partir de l'équation 4.4. Les variables λ_{ir} sont déterminées à partir des valeurs de e_{ir} dans l'équation 4.5. En effet, le taux de visite d'une classe de flux à une station spécifique provient de la probabilité d'arrivée de paquets depuis l'extérieur et depuis les autres stations.

$$\lambda_i = \sum_{r=1}^R \lambda_{ir} \quad (4.4)$$

$$e_{ir} = p_{0,ir} + \sum_{j=1}^M e_{jr} \times p_{ji,r} \quad (4.5)$$

L'équation 4.4 renvoie à la résolution d'un système d'équations. Notons $E_r = \|e_{ir}\|$ et $A_r = \|p_{0,ir}\|$ respectivement les vecteurs des taux de visites et d'arrivées des paquets de classe r . Ainsi, nous obtenons le système suivant :

$$E_r = A_r + P_r' E_r \quad (4.6)$$

TABLE 4.1 – Les paramètres du modèle analytique

Notation	Paramètre
R	Nombre de flux de trafic distincts dans le réseau
N_a	Nombre de flux de trafic d'applications analysées
λ_r	Taux d'arrivées des paquets dans le réseau pour chaque flux de trafic r
M	Nombre de stations qui représentent les liens du réseau
N	Nombre de sommets qui représentent les PIDs
LC_i	Capacité de transmission du lien i
$\mu_i = \mu \times i$	le taux de service au niveau du lien i
λ_{ir}	Taux d'arrivée des paquets du service r au niveau de la station i
$p_{0,ir}$	Probabilité d'arrivée de paquets du service r au niveau de la station i
$p_{ir,0}$	Probabilité de départ de paquets du service r au niveau de la station i
$p_{ij,r}$	Probabilité d'acheminement de paquets du service i des stations i à j
e_{ir}	Taux de visite des paquets du service r à la station i
P_r	Matrice $\ M \times M + 1\ $ de transition des paquets du service r entre les stations

L'équation 4.6 nous permet d'établir l'équation suivante :

$$E_r = (I - P_r)^{-1} A_r \quad (4.7)$$

La matrice $(I - P_R)$ est une matrice stochastique avec un rayon spectral inférieur à un dans notre modèle, car il s'agit d'un modèle de réseau ouvert avec des départs. La résolution du système dans l'équation fournit les valeurs taux de visite à chaque station.

4.1.3.4 Les mesures à l'état d'équilibre

Après la détermination des conditions de stabilité du système, nous pouvons établir l'équation sous forme de produit qui caractérise l'état d'équilibre du réseau de files d'attente BCMP. Cependant, nous ne présenterons pas dans ce document l'équation et les étapes pour sa détermination. Ces travaux mathématiques sont largement expliqués dans la littérature comme dans [BCMP75, Bay00]. En effet, notre intérêt se porte sur les mesures de fonctionnement et de performances fournies par le réseau BCM. En particulier, nous sommes intéressés par des indicateurs comme :

- Les taux d'arrivée λ_{ir} des paquets de chaque classe r dans chaque station i
- Le taux d'utilisation ρ_{ir} de chaque station i par chaque classe r (voir équation 4.8)
- Le nombre moyen de paquets X_{ir} de chaque classe de r à chaque station i (voir équation 4.9)
- La durée moyenne de temps T_{ir} passée par un paquet de classe r à la station i à partir de la formule de Little (voir équation 4.10)

$$\rho_{ir} = \lambda_r \times \frac{e_{ir}}{\mu_i} \quad (4.8)$$

$$X_{ir} = \frac{\rho_{ir}}{1 - \rho_i} \quad (4.9)$$

$$T_{ir} = \frac{X_{ir}}{\lambda_{ir}} \quad (4.10)$$

4.1.3.5 Les métriques de performance

Les résultats obtenus à partir du modèle BCMP nous permettent de calculer des métriques de performances pour évaluer à la fois les réseaux de transport et les applications qui les exploitent. A titre d'exemple, l'équation 4.4 fournit le débit de trafic en paquets/s sur chaque lien du réseau. En connaissant les capacités de transport des liens réseaux, nous pouvons calculer le coût d'acheminement du trafic LC_i sur chacun d'eux (équation 4.11) et en conséquence le coût d'utilisation global du réseau $Netcost$ (équation 4.12).

$$LC_i = a \times \exp\left(b \times \frac{\lambda_i}{LCAP_i - \lambda_i}\right) \quad (4.11)$$

$$Netcost = \sum_{i=1}^M LC_i \quad (4.12)$$

Nous pouvons aussi déterminer la durée moyenne d'attente dans chaque station (équation 4.13) et en dériver la latence moyenne sur chaque segment du réseau entre un PID source i et un PID destination j (équation 4.14).

$$T_i = \sum_{r=1}^R \frac{\lambda_{ir} \times T_{ir}}{\lambda_i} \quad (4.13)$$

$$PTT_{ij} = \sum_{u \in \pi_{ij}} T_u \quad (4.14)$$

Les temps de traitement et de propagation peuvent être utilisés aussi dans le calcul de la métrique PTT s'ils sont considérés non négligeables. Ces temps de latence par segment nous permettent aussi d'en dériver un indicateur global de la qualité de transport de données pour chaque application r en calculant une latence moyenne $APTT_r$ sur les tous les segments empruntés par l'application (équation 4.15).

$$APTT_r = \sum_{i=1}^N \sum_{j=1}^N \frac{NC_{ij}^r \times PTT_{ij}}{NC^r} \quad (4.15)$$

Dans l'équation précédente, NC_{ij}^r représente le nombre de clients de chaque application r qui utilisent le segment π_{ij} du réseau alors que NC^r représente le nombre total des clients de l'application.

D'autres métriques et indicateurs de performances sont aussi soit directement disponibles soit calculables à partir des paramètres modèles. Nous citons par exemple, le nombre de liens utilisés par chaque application ou bien la latence moyenne du réseau.

4.1.4 La configuration du modèle analytique pour les simulations

L'outil de modélisation que nous avons présenté dans les sections précédentes nous a permis de réaliser des simulations pour mesurer les performances des techniques d'optimisations que nous avons proposées dans les chapitres 2 et 3. L'objectif que nous visions était de comparer les techniques que nous avons développées à la fois entre elles mais aussi avec des techniques usuelles à travers la modélisation de l'exécution de services et applications différentes utilisant un réseau que nous maîtrisons.

4.1.4.1 Implémentation du modèle

Le modèle analytique a été développé sur Matlab. Cet environnement fournit un panel d'outils très avancés pour les traitements mathématiques et matriciels, en plus des capacités de programmation. Le programme a été structuré de manière conforme à l'exécution discrète de la simulation.

Les étapes de la simulation Chaque simulation répond à un schéma d'exécution présenté dans la figure . Le modèle du réseau est d'abord initialisé et configuré. Les PIDs et les liens entre eux sont créés. Leurs capacités sont définies. Le routage est déterminé entre les PIDs source et destination. Durant la seconde étape, le trafic réseau est initialisé aléatoirement. Ce trafic représente le trafic dans le réseau qui n'est pas dans notre champ d'action. La troisième étape permet ensuite d'initialiser les paramètres d'exécution de la simulation comme le nombre de répétition mais aussi d'autres paramètres applicatifs comme les débits de transmission des données.

L'étape suivante est un ensemble de sous-étapes répétitives. En effet, chaque itération correspond à un pas discret dans le temps. Ces itérations modélisent l'évolution dans le temps de la simulation. Elles commencent par une variation aléatoire du trafic dans le réseau pour modéliser l'évolution du trafic dans le temps. Ensuite, le module de collecte des informations réseaux est exécuté à la fois pour déterminer les performances du réseau mais aussi pour disposer des données qui seront transmises aux applications. Ces informations seront disposées dans les formats adéquats : carte de réseau, matrice de coût, matrice de contraintes, matrice de proportions. Ces données sont exploitées ensuite par les applications concernées pour appliquer les différentes techniques d'optimisation à comparer. Le même contexte d'exécution est utilisé pour assurer une base commune pour la comparaison. Ainsi le nombre de clients, de serveurs, leurs capacités ou encore le débit des données sont similaires. La dernière sous-étape de chaque itération est la mesure des performances des différentes techniques.

4.2 Évaluation de l'approche passive de collaboration

Nous présenterons dans cette section les résultats des simulations que nous avons effectué pour évaluer les performances de l'approche passive de collaboration.

4.2.1 Présentation des paramètres

Nous avons déjà évoqué dans la section précédente la génération aléatoire de certains paramètres comme le nombre de clients dans le réseau. Certains autres paramètres comme la topologie des réseaux modélisés sont déterministes. Nous nous sommes appuyés sur des bases d'informations à la fois fournies par les opérateurs ou bien par des organismes comme CAIDA, "*The Cooperative Association for Internet Data Analysis*" [CAI], qui recueillent des informations sur les réseaux de transport. Nous avons utilisé ainsi des informations sur la topology du coeur de réseau de l'opérateur *AT&T* pour modéliser le réseau dans la simulation. Chaque PoP de la topologie, représentant un sommet du graphe (total de 46), est considéré comme un PID dans la carte réseau qui sera fournie aux services applicatifs à travers le protocole ALTO.

Pour simuler des conditions d'exploitation différentes, nous avons fait varié plusieurs paramètres dont l'emplacement des serveurs dans le réseau, l'emplacement et le nombre de clients et la charge du réseau initial sur les liens. Ainsi pour chaque simulation, trois serveurs sont situés aux niveaux de PoPs de grande incidence, c'est-à-dire, des

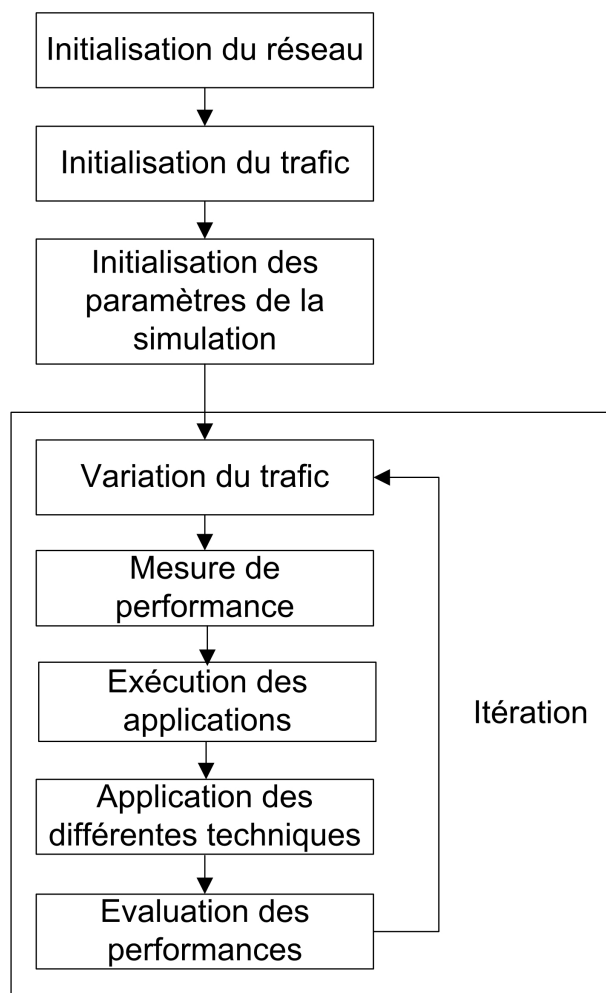


FIGURE 4.2 – Les étapes d’executions des simulations analytiques

PoPs importants inter-connectés avec un nombre d’autres PoPs conséquent. Ce choix se justifie dans la réalité par le fait que les noeuds de distribution seront déployés en priorité dans les zones de forte densité de population pour minimiser la distance vis-à-vis du plus grand nombre possible d’utilisateurs. Ces PoPs sont généralement situés dans les grandes agglomérations et s’inter-connectent avec divers autres PoPs.

Les conditions de trafic sont elles aussi variées. Nous avons choisi dans la simulation de nous contenter à deux niveaux d’utilisation moyenne du réseau ; *20%* et *40%* pour simuler des périodes d’utilisation normale et d’utilisation plus sévère. Cette moyenne s’établit sur l’ensemble des liens du réseau, qui quant à eux sont soumis à des taux d’utilisation totalement aléatoires. Les simulations sont ensuite répétées une centaine de fois pour chaque niveau d’utilisation du réseau et pour chaque emplacement particulier des noeuds de distribution dans le réseau, avec une génération d’un nombre aléatoire de clients

à chaque fois. Par ailleurs, comme nous l'avons évoqué précédemment, pour chaque simulation, un certain nombre d'itérations sont exécutées. Les principaux paramètres de la simulation sont exposés dans le tableau 4.2.

TABLE 4.2 – Principaux paramètres des simulations

Paramètre	Utilisation à 20%	Utilisation à 40%
Niveau moyen initial d'utilisation réseau	20.16 %	40.42 %
Nombre moyen de clients	2915	5854
Niveau moyen total d'utilisation réseau	24.86 %	48.79%
Nombre d'emplacements différents des serveurs	4	4
Débit applicatif	[0.9, 1.5] Mbps	[0.9, 1.5] Mbps
Nombre de simulations	400	400

4.2.2 Résultats des simulations

Nous avons évalué les résultats de différentes techniques de sélection d'un noeud de distribution pour chaque client pendant les simulations effectuées :

1. Technique de sélection aléatoire d'un noeud
2. Technique de sélection du noeud le plus proche en utilisant une métrique distance dans la matrice de coût. La distance est calculée à partir du nombre de liens traversés entre le client et le serveur.
3. Technique de sélection d'un noeud à partir de la métrique latence dans la matrice de coût.
4. Technique de sélection d'un noeud à partir de la métrique coût de transport. Ce coût est calculé à partir des ressources disponibles sur les différents chemins envisageables entre le client et tous les serveurs présents dans le réseau.
5. Technique de sélection d'un noeud de distribution par l'opérateur réseau. L'application soumet une requête à l'opérateur pour choisir un noeud adéquat parmi une liste de serveur pour un client particulier. Cette technique se base aussi sur la métrique coût de transport. La différence avec la technique précédente est que les coûts de transport sont mis à jour en temps réel alors que la matrice de coût envoyée aux applications est mise à jour après une certaine période de temps.

Nous récupérons ensuite différentes métriques des simulations pour évaluer les performances de transport des flux de données, celles du réseau et celles des services applicatifs. Avant d'exposer un peu plus en détail la teneur de ces résultats, nous proposons d'étudier la redirection des clients dans un réseau composé uniquement de 19 PIDs. Ce nombre limité a été choisi uniquement pour nous faciliter l'analyse des redirections des clients. Les simulations par contre ont porté sur des réseaux plus étendus. Le graph

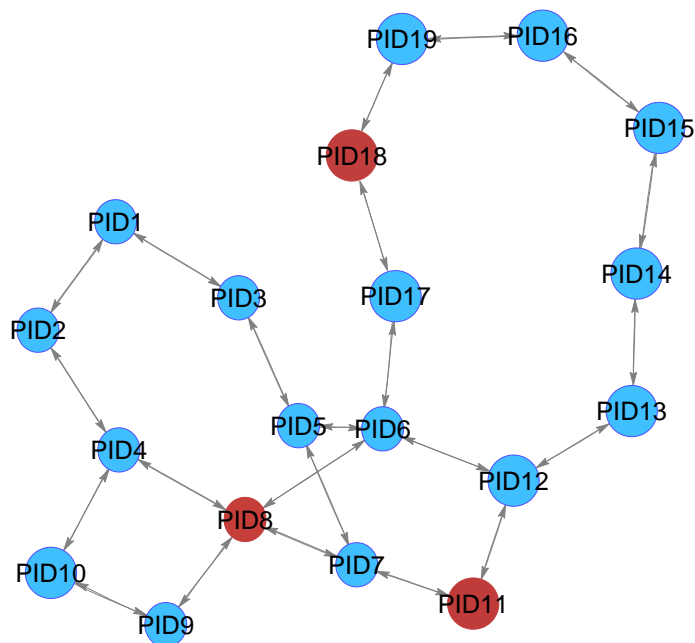


FIGURE 4.3 – Graph du réseau RNP

(voir figure 4.3) modélise le réseau "Rede Nacional de Ensino e Pesquisa" (RNP), qui fournit les informations de topologie et de capacités de transport [dEeP13]. Dans cet exemple, les serveurs ont été placés dans les PIDs 8, 11 et 18 coloriés en rouge.

Redirection des clients Dans la figure 4.3, nous présentons le nombre de redirection de clients vers chaque serveur effectuées par chaque technique. La sélection aléatoire de serveur résulte en l'utilisation de tous les serveurs par la quasi totalité des PIDs, c-à-d que les clients dans chaque PID utilisent le lot entier de serveurs disponibles. Cette technique est clairement inefficace. La métrique distance est au contraire exclusive entre PIDs. En effet, même quand il y a deux PIDs à égale distance d'un autre, un seul est retenu. Au final, les clients dans chaque PID sont servis par un seul et unique serveur dans le PID le plus proche. Finalement, les trois autres métriques qui dépendent des conditions de transport dans le réseau évolue en fonction du temps. Nous pouvons apercevoir que les redirections des clients ne sont exclusives ni aléatoires entre les PIDs

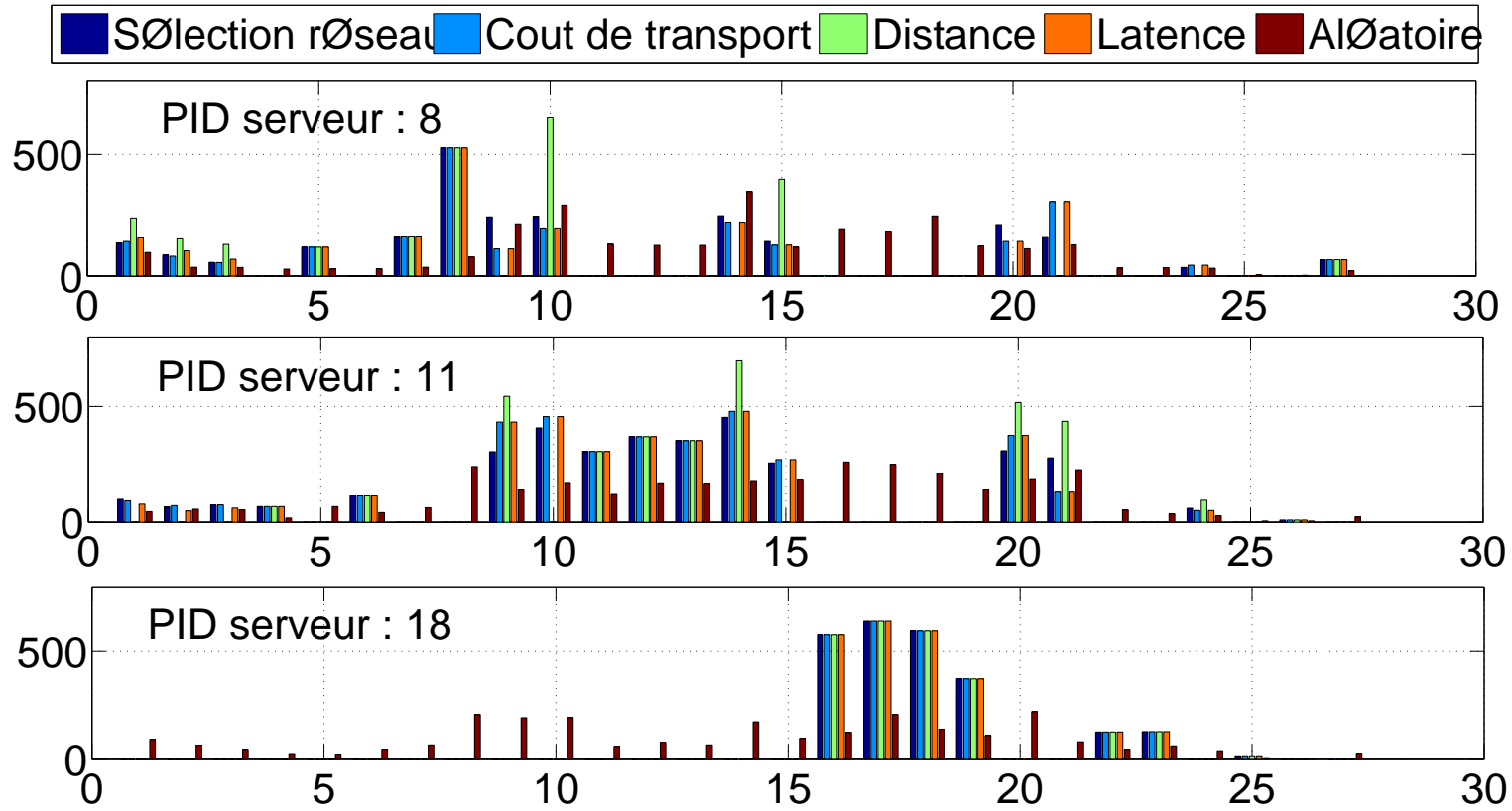


FIGURE 4.4 – Graph du réseau RNP

Gain en utilisation moyenne du réseau Les résultats obtenus démontrent que les techniques de sélection qui sont basées sur des métriques du réseau permettent de réaliser un gain non négligeable en termes d'utilisation des ressources du réseau par rapport à une redirection aléatoire des clients vers les noeuds de distribution des contenus. Sur un échantillon de simulations exposé dans la figure 4.5, nous pouvons constater que le gain réalise oscille entre 41% et 52% en moyenne en fonction des techniques. Dans ce cadre, c'est bien la technique de sélection basée sur la métrique distance qui réalise les meilleurs gains. Ce résultat s'explique par le fait que cette technique engendre l'utilisation du plus petit nombre de liens possibles comme chaque client est redirigé vers le serveur le plus proche au sens nombre de liens réseaux utilisés. En conséquence, l'utilisation du réseau est minimale et le gain est maximal. Notons tout de même, qu'il n'y a pas une très grande différence par rapport aux autres techniques analysées.

Distribution du taux d'utilisation des liens Le taux d'utilisation des liens du réseau est un aspect important à considérer pour évaluer les bénéfices apportés par chaque technique. En effet, la condensation du trafic sur un nombre limité de liens est susceptible d'engendrer des problèmes de congestion à cause d'une charge de trafic mal répartie dans le réseau. Nous constatons à travers les résultats que nous avons collectés que l'utilisation de la métrique coût de transport permet de réaliser la meilleure répartition de charge dans le réseau. En comparant la distribution des taux d'utilisation des liens par plage de 10% dans la figure 4.6, le pourcentage de liens dont le taux d'utilisation est supérieur à 90% est respectivement de 20%, 10%, 8% et inférieur à 2% pour les techniques basées sur une sélection aléatoire, sur la métrique distance, sur la métrique latence et sur la métrique coût de transport. Ces pourcentages se rééquilibrent ensuite pour des taux d'utilisation supérieur à 70% autour de 30% pour les techniques qui se basent sur des informations réseaux. Par contre la sélection aléatoire présente un pourcentage de liens plus important (43%). Nous voyons bien en conséquence le bénéfice apporté par la technique basée sur la métrique coût de transport que nous proposons. Elle permet de réaliser une meilleure répartition de charge dans le réseau et évite des taux d'utilisation élevés sur un grand nombre de liens. En conséquence, en contre-partie d'un gain inférieur en utilisation des ressources du réseau, la maîtrise de la charge est mieux assurée.

Ces constatations moyennées sur l'ensemble des résultats obtenus des simulations sont confirmées par l'analyse de l'utilisation des liens du réseau obtenues à la fin de chaque simulation. Ces résultats démontrent aussi que l'exploitation d'informations de transport dans la problématique de gestion de trafic apporte des avantages indéniables. Si les performances obtenues pour un niveau de trafic initial à 20% semblent modérées, l'intérêt est par contre évident pour une utilisation moyenne des ressources du réseau aux environs de 40%. En effet, à 20%, le réseau dispose encore de 80% de ces ressources disponibles d'où une meilleure capacité pour faire face aux flux de trafic. En conséquence, les taux d'utilisation des liens restent tout de même dans des proportions non alarmantes (figure 4.7. Pour un niveau de trafic moyen de 40%, le nombre de liens dont le taux d'utilisation dépasse 80% est plus important (figure 4.8). Par ailleurs, toutes les

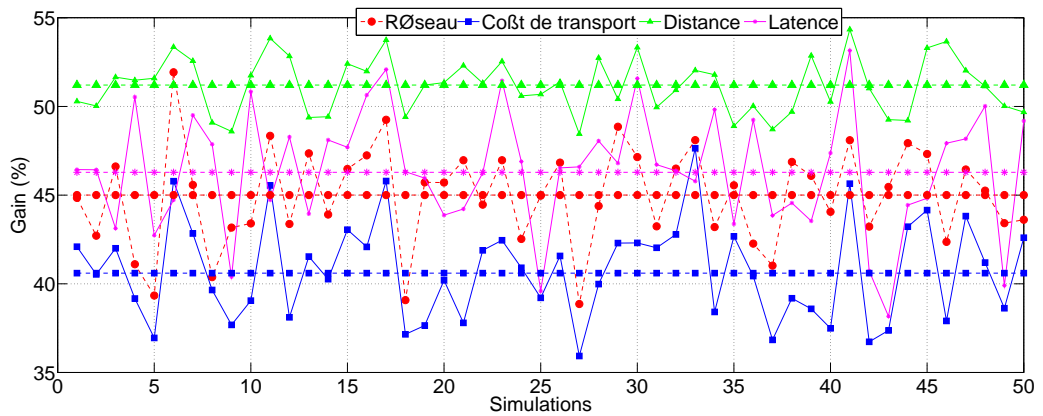


FIGURE 4.5 – Le gain réalisé en termes d'utilisation moyenne du réseau des différentes techniques de redirection par rapport à une sélection aléatoire

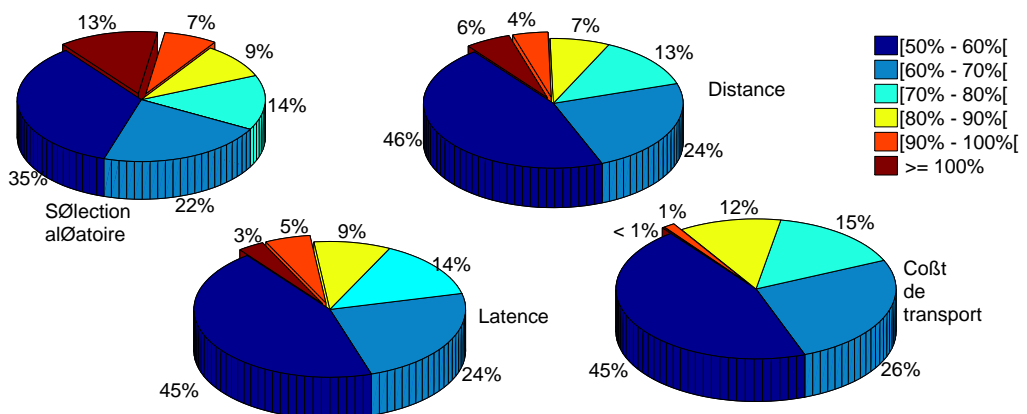


FIGURE 4.6 – La distribution des taux d'utilisation des liens du réseau des différentes techniques de redirection

approches à l'exception de la sélection qui est opérée par l'opérateur réseau, présente des liens utilisés à plus de 80%. En effet, même une matrice de coût fournissant une métrique de coût de transport ne peut éviter ces taux d'utilisation élevés si la fréquence de mise à jour des informations n'est pas élevée. Rappelons que les coûts de transport augmentent exponentiellement avec les taux d'utilisation des liens avec interdiction de dépassement de 80% de la capacité dans notre modèle.

La variance de l'utilisation de liens réseaux offre aussi une excellente analyse de l'apport de chaque technique en termes d'utilisation des ressources du réseau. Elle caractérise en effet la dispersion des taux d'utilisation des liens. Nous constatons encore une fois à travers nos résultats que la métrique de coût de transport dont nous proposons l'utilisation dans la matrice de coût présente la variance la plus faible utilisation par rapport aux autres approches. Les moyennes d'utilisation sont équivalentes pour les techniques qui exploitent des informations réseaux (figure 4.9). Cette variance plus faible indique une faible dispersion des taux d'utilisation des liens. Pour des moyennes équivalentes, elle assure des pics d'utilisation moins fréquents que les autres technique ce qui résulte en une meilleure répartition de charge au sein du réseau.

Latence moyenne réduite Les services applicatifs tirent un bénéfice aussi de l'approche de collaboration passive que nous proposons. En effet, ils profitent d'une meilleure disponibilité de ressources dans le réseau pour leurs flux de données par le biais d'une meilleure répartition de charge. Les utilisateurs sont aussi bénéficiaires des lots d'amélioration à travers une latence moyenne réduite dans le réseau. Les résultats que nous avons obtenus démontrent que la latence moyenne sur l'ensemble des chemins du réseau est plus faible en exploitant les coûts de transport (figure 4.10). En effet, lorsque la charge de trafic généré est mieux répartie sur les différents équipements du réseau, leurs temps de réponse sont plus faibles. En conséquence, même si certains chemins peuvent être un peu plus longs, les temps de parcours de bout en bout pourraient être réduits grâce à cette meilleure disponibilité du réseau. Notons aussi que la probabilité de perte de paquets se trouve aussi réduite et que les risques de congestion sont plus limités. Tous ces facteurs contribuent à fournir une meilleure qualité d'expérience pour les utilisateurs.

4.2.3 Synthèse et conclusions

Compte tenu des résultats que nous avons obtenus à partir des simulations réalisées, nous pouvons effectuer un classement des performances des différentes techniques évaluées pour l'approche de collaboration passive. Dans ce sens, ce sont les technique qui exploitent les coûts de transport qui fournissent les meilleurs résultats par rapport aux autres en termes de répartition de charge sans pénaliser les performances du réseau mesurées en termes de latence moyenne. La métrique latence peut fournir des indications approximatives par rapport à l'état d'utilisation du réseau. Ainsi, elle peut aider à éviter la concentration du trafic sur des liens présentant des taux d'utilisation déjà élevés. Cependant, au-delà des résultats de performance, certaines approches ne sont

4.2. Évaluation de l'approche passive de collaboration

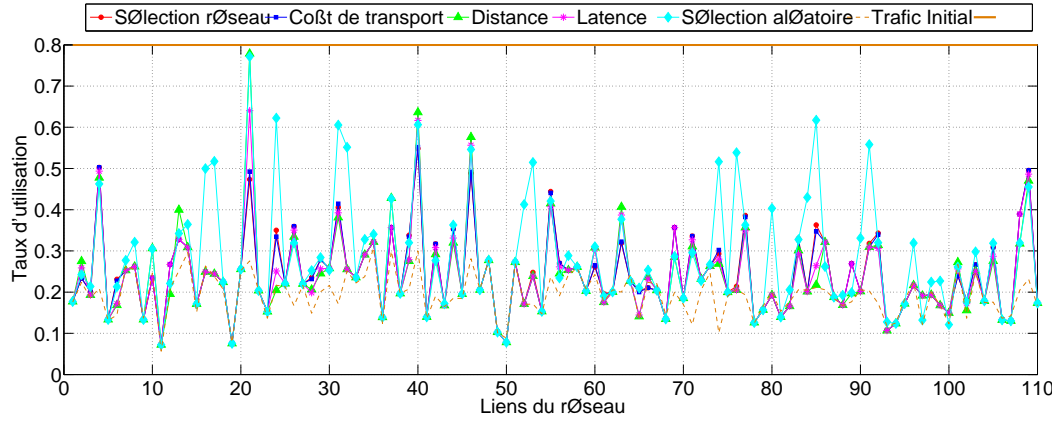


FIGURE 4.7 – Illustration de l'utilisation des liens du réseau durant une simulation pour un trafic initial à 40%

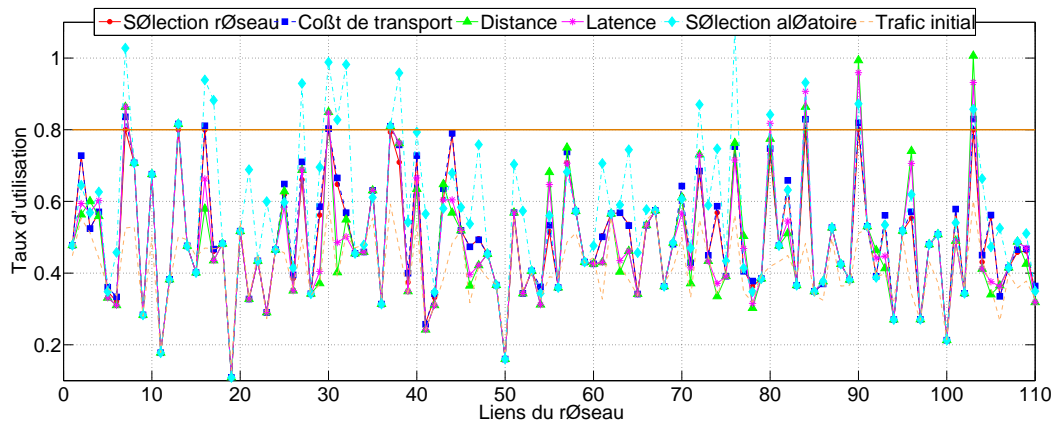


FIGURE 4.8 – Illustration de l'utilisation des liens du réseau durant une simulation pour un trafic initial à 40%

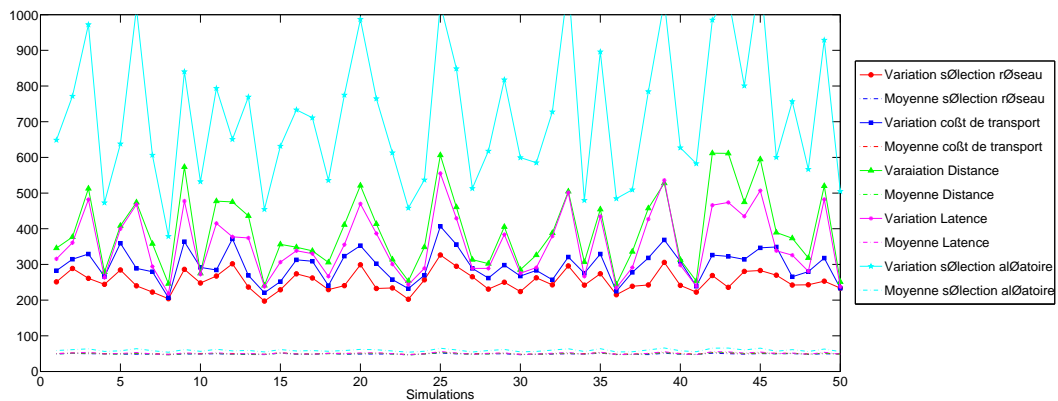


FIGURE 4.9 – Illustration de la moyenne et variance des taux d'utilisation des liens

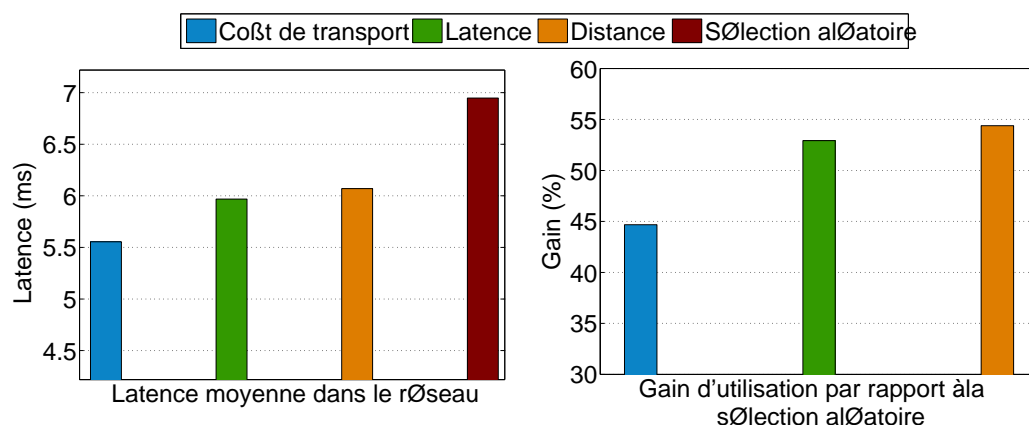


FIGURE 4.10 – Illustration des gains réalisés en utilisation et en latence des différentes techniques

pas crédibles pour une réelle exploitation. Ainsi la technique de sélection opérée par le réseau présente un inconvénient majeur. Le mécanisme de requête/réponse résulte en une énorme charge sur les infrastructures dédiées et un délai supplémentaire pour la redirection des clients. La technique de sélection aléatoire quant à elle fournit des résultats très médiocres. En revanche, la métrique distance peut fournir des performances similaires aux coûts de transport en particulier dans des contextes favorables, c'est-à-dire quand les taux d'utilisation du réseau sont faibles.

Il est également important de signaler l'incidence de la fréquence de mise à jour de la matrice de coût sur les performances des techniques dont qui se basent sur des informations dynamiques comme le coût de transport. En effet, lorsque le réseau est soumis à de sévères conditions, des mises à jour fréquentes des coûts de transport conduisent à une amélioration des performances. Ces aspects doivent en conséquence être pris en considération par les opérateurs de réseaux et acteurs applicatifs au moment de se fier à l'approche de collaboration passive pour améliorer les performances de leurs systèmes.

4.3 Évaluation de l'approche active de collaboration

Nous présenterons dans cette section les résultats des simulations réalisées pour évaluer les performances de l'approche active de collaboration.

4.3.1 Présentation des paramètres

La configuration des simulations dans cette section héritent des paramètres utilisés pour l'évaluation de l'approche passive. A ce titre, le modèle analytique, les données topologiques, la génération aléatoire des valeurs de certains paramètres sont identiques. Le modèle d'exécution des simulations est lui aussi similaire. De nouveaux paramètres sont spécifiques cependant au modèle de collaboration active. Il s'agit tout particulièrement des informations renseignées par les services applicatifs. Nous avons évoqué lors de la présentation de cette approche que les applications fournissent une estimation du trafic qu'elles s'appêtent à générer dans le réseau. Dans notre modèle, cette estimation portent sur le nombre de clients qu'elles pensent servir sur une plage de temps jusqu'à la mise à jour des informations. En conséquence, nous avons introduit un taux d'erreur dans notre modèle pour compenser l'écart en pratique entre le nombre réel de clients qui seront servis et le nombre estimé. Ce taux d'erreur varie aléatoirement entre $\pm[20\%, 30\%]$ pour chaque simulation. La capacité des serveurs est un autre facteur déterminant dans la résolution du problème et renseigné par les services applicatifs. Nous détaillons l'ensemble des principaux paramètres dans le tableau 4.3.

TABLE 4.3 – Principaux paramètres des simulations

Paramètre	Utilisation à 20%	Utilisation à 40%
Niveau moyen initial d'utilisation réseau	20.73 %	40.78 %
Taux d'erreur des estimations	$\pm[20\%, 30\%]$	$\pm[20\%, 30\%]$
Nombre moyen réel de clients dans le réseau	5244	7939
Nombre moyen estimé de clients dans le réseau	5238	7946
Niveau moyen total d'utilisation réseau	$[22.88\%, 24.93\%]$ %	$[44.01\%, 47.11\%]$
Nombre d'emplacements différents des serveurs	4	4
Nombre de serveurs	[3, 4]	[3, 4]
Capacité totale des serveurs (en nombre de clients)	10 000	10 000
Débit applicatif	[0.9, 1.5] Mbps	[0.9, 1.5] Mbps
Nombre de simulations	400	400

4.3.2 Résultats des simulations

Pour évaluer les performances de l'approche active de collaboration que nous proposons, nous avons comparé ses résultats avec les résultats de différentes techniques de sélection d'un noeud de distribution :

1. La technique de sélection aléatoire
2. La technique de sélection d'un noeud à partir de la métrique coût de transport.

Nous récupérons ensuite différents résultats des simulations pour évaluer les performances de transport des flux de données, celles du réseau et celles des services applicatifs.

Redirection des clients Nous présentons dans la figure 4.11 une illustration du réseau dorsal européen d'Orange que nous avons modélisé pour nos simulations. Pour un certain d'entre elles, les noeuds de distribution ont été positionnés dans les PIDs de couleur rouge. Nous nous intéressons particulièrement à comparer les redirections de clients définies par les deux approches de collaboration (voir la figure 4.12) car la nature de la sélection aléatoire permet de réaliser une bonne répartition de la charge sur les noeuds de distribution mais elle résulte en l'utilisation de tous les serveurs par la quasi totalité des PIDs, ce qui est loin d'être performant du point de vue transport et service. La première constatation que nous pouvons relever est l'exclusivité d'un serveur pour chaque PID pour la collaboration passive. Du fait de l'utilisation d'une matrice de coût, un seul chemin réseau est retenu ce qui aboutit à l'utilisation du même serveur par tous les clients appartenant au même PID. Ainsi, les PIDs sont distinctement répartis sur les serveurs, qui sont généralement à proximité. Quand à la technique de collaboration active, les clients du même PID peuvent être servis par différents noeuds de distribution. Nous constatons ainsi que le nombre de PIDs servis par chaque noeud est supérieur pour l'approche active par rapport à l'approche passive. La performance de la technique que nous proposons repose au final énormément sur les proportions qui définiront le nombre de clients à rediriger vers chaque serveur dans les PIDs du réseau.

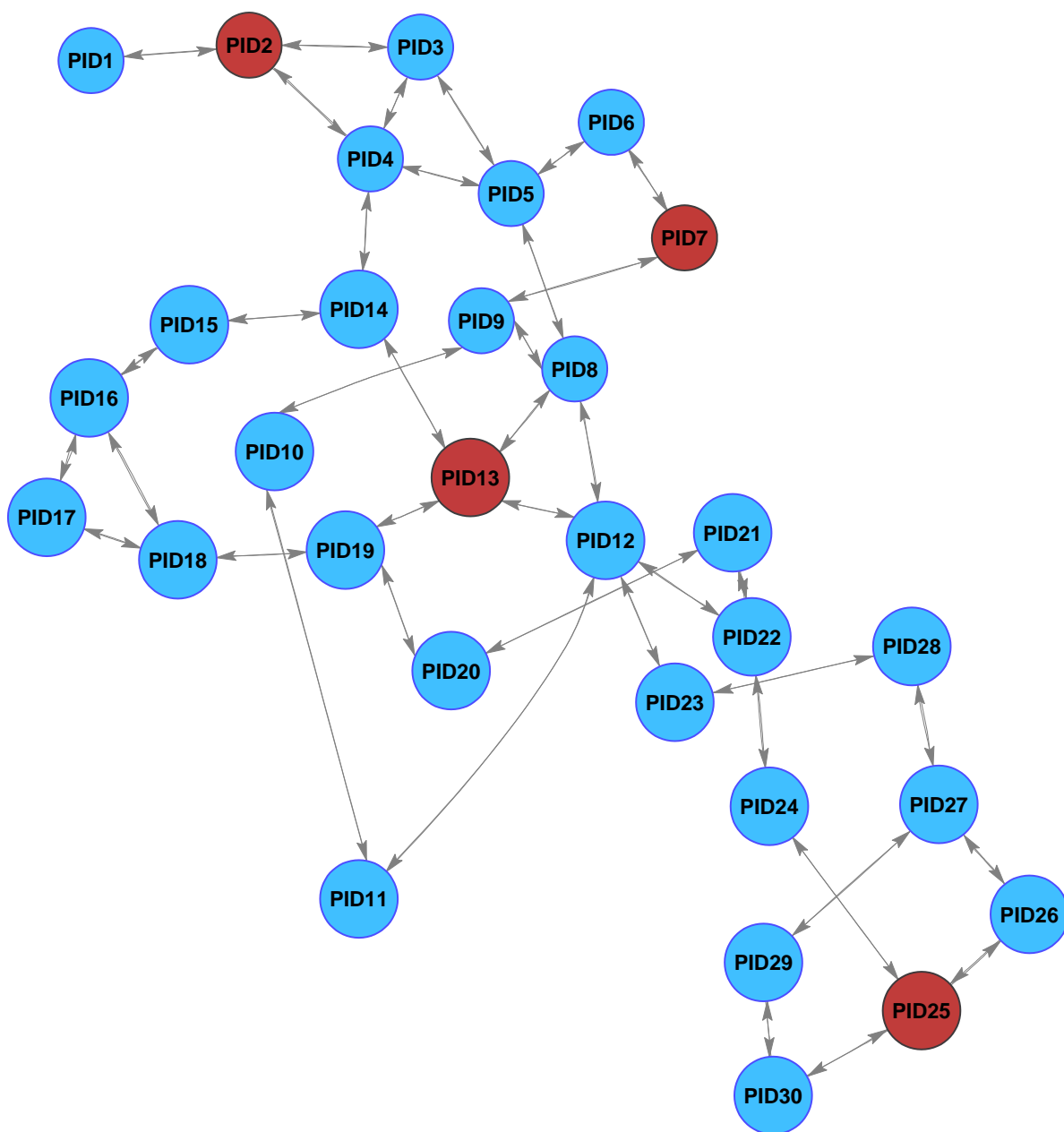


FIGURE 4.11 – Graph du coeur de réseau Orange Europe

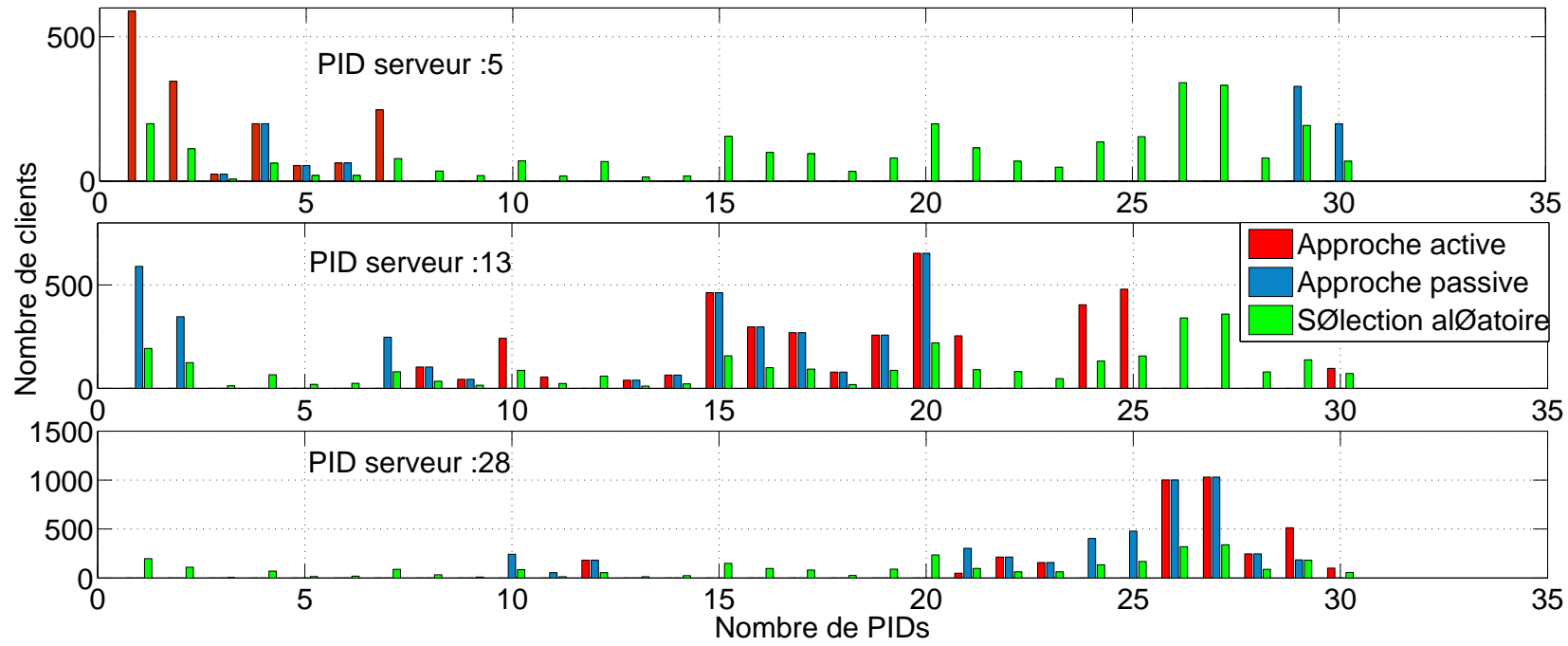


FIGURE 4.12 – Illustration du nombre de clients redirigés vers les différents serveurs inter-connectés avec le réseau

Taux d'utilisation des liens Les taux d'utilisations des liens du réseau est un indicateur d'une bonne distribution de charge. Il s'agit d'un aspect important à prendre en considération pour exploiter pleinement et de manière efficace l'ensemble des ressources disponibles dans le réseau. Nous présentons les taux d'utilisation des liens au cours une simulation unique dans la figure 4.13 (car il n'est pas possible de faire une représentation de ces valeurs par rapport à l'ensemble des simulations au regard de certains paramètres aléatoires qui varient). Nous remarquons dans cette figure quelques pics d'utilisation qui correspondent à des liens sur-utilisés. L'approche active de collaboration est par ailleurs la seule qui arrive à respecter la limite d'utilisation des liens que nous avons fixé à 80%. L'intérêt qui peut sembler léger dans ce cas particulier est non négligeable. En effet, un lien très utilisé dans le réseau est généralement un lien qui appartient à plusieurs chemins liant différents PIDs du réseau. Si un tel lien est congestionné, toutes les entités applicatives dans tous les PIDs qui exploitent des chemins réseaux incluant ce lien seront affectées. Pour ces raisons, nous partons dans notre approche du constat que chaque lien est important mais aussi que ceux qui sont les plus utilisés sont généralement les plus importants. En arrivant à limiter l'utilisation des liens et à les contenir sous certains de sécurité, nous assurons une meilleure résilience dans le réseau aux bénéfiques des applications et de leurs utilisateurs.

Distribution des taux d'utilisation des liens Nous avons comparé la distribution des taux d'utilisation des liens pour les différentes approches par rapport à l'ensemble des simulations que nous avons réalisées. Ces valeurs nous permettront de juger de l'efficacité de chaque technique quand il s'agit d'assurer une bonne répartition de la charge dans le réseau. Nous constatons à travers les résultats que nous avons obtenus que l'approche active de collaboration assure le maintien de l'utilisation des liens en dessous de 80% (figure 4.14). Il y a un report de charge sur d'autres liens et c'est la raison pour laquelle la plage d'utilisation entre 50% et 60% est de 79% alors qu'elle est beaucoup plus faible pour les autres approches, 44% pour la sélection aléatoire et 48% pour la technique de sélection basée sur les coûts de transport. Le bénéfice réalisé est donc non négligeable sur plusieurs niveaux. En effet, il permet d'assurer une meilleure utilisation des ressources du réseau, de supporter une charge supérieure par rapport aux techniques et comme nous allons voir, d'offrir une meilleure qualité de service.

Ces constatations sont corroborées par ailleurs par les valeurs de variation des taux d'utilisation des liens illustrées dans la figure 4.15. L'approche de collaboration active présente la dispersion la plus faible pour les taux d'utilisation ce qui traduit une meilleure répartition de la charge par rapport aux autres techniques.

Gain en utilisation moyenne du réseau Par rapport à la sélection aléatoire, les gains qui peuvent être atteints en termes d'utilisation moyenne du réseau sont très intéressants. La moyenne des gains sur l'ensemble des simulations que nous avons effectuées a été de l'ordre de 45% pour la collaboration active et de 52% pour la collaboration passive (figure 4.16). Les résultats obtenus démontrent que la collaboration entre les acteurs du transport et les acteurs du service est une option concrète dans l'optimisation de leurs environnements. Et si le gain réalisé par la collaboration passive est plus

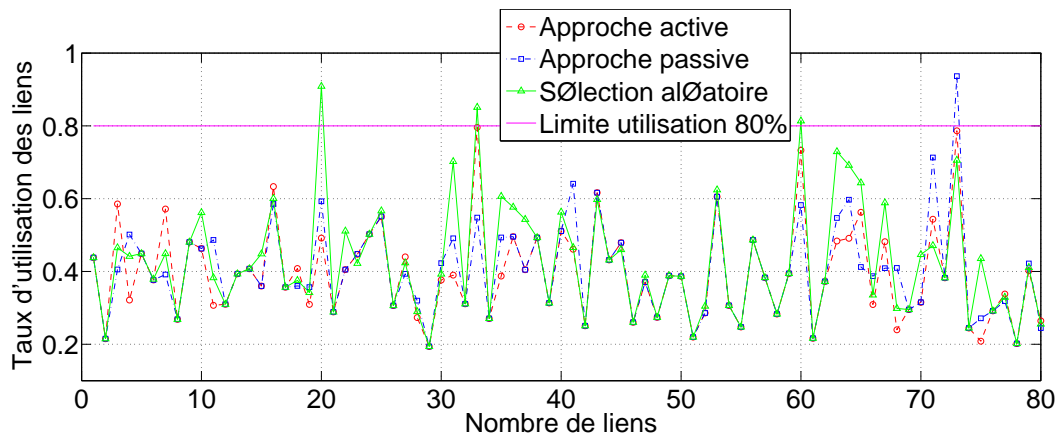


FIGURE 4.13 – Le gain réalisé en termes d'utilisation moyenne du réseau des différentes techniques de redirection par rapport à une sélection aléatoire

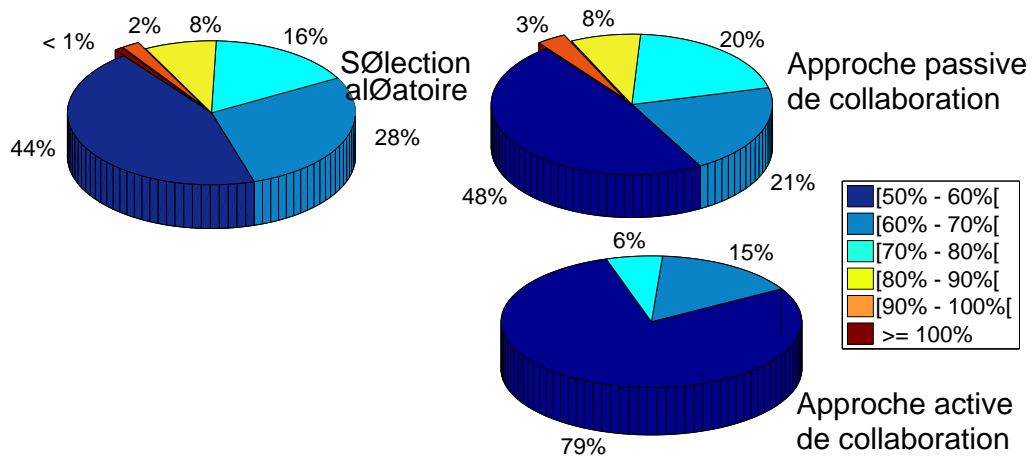


FIGURE 4.14 – La distribution des taux d'utilisation des liens du réseau des différentes techniques de redirection

important, il se paye aux prix d'une moins bonne répartition de charge mais aussi de performances moins intéressantes comme nous allons le voir. Ce gain plus important s'explique par contre par le fait que la collaboration active pousse encore plus loin la possibilité d'utiliser des chemins réseaux plus longs. Le facteur déterminant entre les deux techniques reste les taux d'utilisation des liens au départ du processus et l'arbitre devient leur coût d'utilisation pour l'acheminement du trafic. En conséquence, l'utilisation d'un plus grand nombre de liens par l'approche active assure une meilleure gestion des ressources au détriment d'un gain global légèrement plus faible.

Latence moyenne réduite Nous avons constaté dans les résultats de nos simulations que la latence moyenne dans le réseau est meilleure pour l'approche de collaboration active. La bonne gestion de l'ensemble des ressources du réseau assure une meilleure disponibilité pour l'acheminement du trafic dont tire profit directement les services applicatifs et leurs utilisateurs (figure 4.16).

4.3.3 Synthèse et conclusions

Les résultats fournis par les simulations réalisées nous confortent dans notre théorie de développement d'une nouvelle approche inter-couches qui permet aux différents acteurs de collaborer, mettre à disposition leurs informations et ressources en vue d'assurer une optimisation conjointe de leurs performances. Nous avons bien constaté des gains dans le réseau qui affectent directement et indirectement le trafic des services applicatifs.

L'apport indéniable de la collaboration active n'est pas uniquement dédié au rattrapage des lacunes inhérentes à l'approche de collaboration passive mais surtout à l'amélioration des performances du processus global. Bien que les informations soient mise à jour à des instants discrets dans le temps pour les deux approches, la collaboration active hérite d'un procédé optimal de répartition de charge. Nous avons constaté la différence entre les résultats des deux approches par rapport à ce aspect et leurs répercussions sur l'ensemble des indicateurs de performances que nous avons analysés.

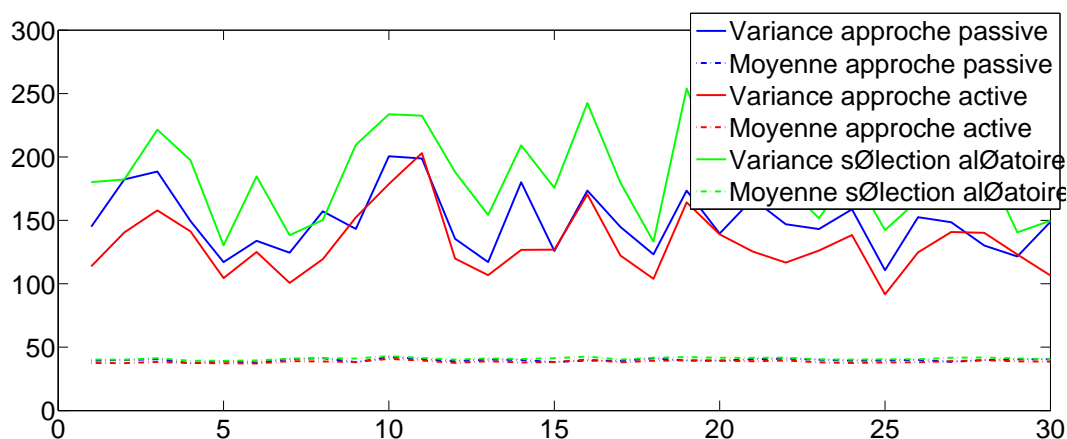


FIGURE 4.15 – Illustration de l'utilisation des liens du rØseau durant une simulation pour un trafic initial à 40%

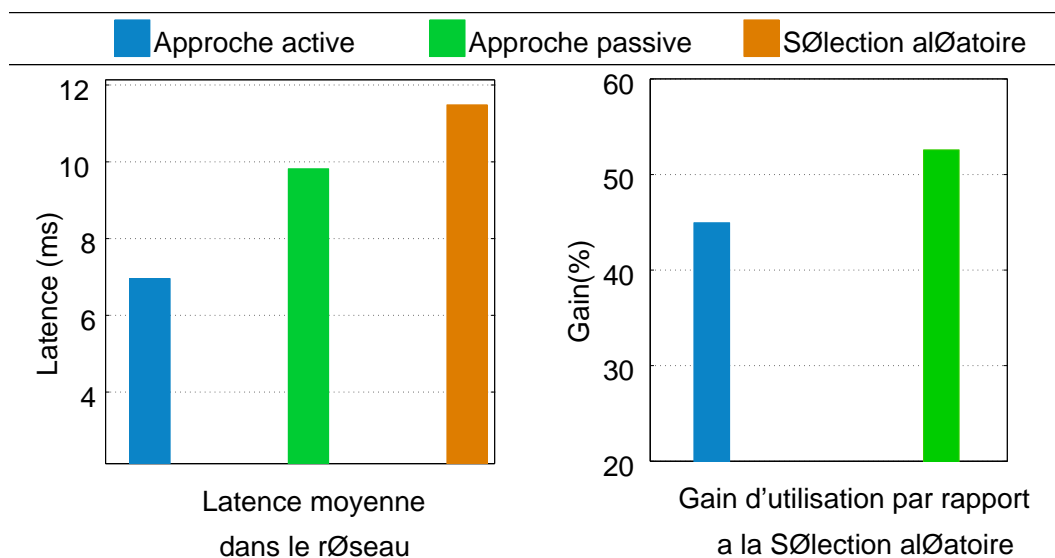


FIGURE 4.16 – Illustration de l'utilisation des liens du rØseau durant une simulation pour un trafic initial à 40%

4.4 Évaluation de l'application de l'approche de collaboration à un système P2P

Nous avons présenté dans la section 3.2 une architecture pair-à-pair optimisée pour distribution de contenus. Nous allons procéder dans cette section à l'évaluation de l'application de l'approche de collaboration dans ce système à travers les performances du réseau et de l'application P2P. Nous présenterons les détails et les résultats des simulations que nous avons effectuées.

4.4.1 Présentation des paramètres de la simulation

Nous avons utilisé le modèle analytique présenté dans la section 4.1 pour modéliser le réseau de transport qui supportera l'application P2P. La différence réside dans la nature du modèle applicatif qui est une architecture P2P spécifique contrairement au modèle client-serveur employé dans les précédentes simulations.

Notre système P2P est caractérisé par des paramètres qui lui sont spécifiques comme la taille de chaque groupe (le nombre de pairs qu'il contient) ou la taille de chaque arbre de groupes. Notre objectif étant d'évaluer le gain en performance que pourrait apporter la mise en oeuvre d'une collaboration inter-couches, nous n'avons pas cherché à optimiser le paramétrage du système P2P mais nous restons conscients qu'il y a là aussi d'intéressantes perspectives. Nous avons procédé en conséquence à des choix de configuration présentés dans la table 4.4. En particulier, notons que nous avons réalisé tout de même les simulations avec trois tailles d'arbres de groupes différentes 5, 10, 15 et dans les mêmes conditions.

TABLE 4.4 – Principaux paramètres des simulations

Paramètre	Utilisation faible	Utilisation moyenne
Nombre de pairs par groupe	9	
Nombre de groupes par arbre	6	
Capacité excédentaire de transmission par groupe	2	
Taux moyen initial d'utilisation réseau	20.35 %	40.34 %
Taux moyen d'utilisation réseau après simulation	22.69 %	42.96 %
Nombre moyen de pairs dans le réseau	3513	4887
Débit applicatif	[0.9, 1.5] Mbps	[0.9, 1.5] Mbps
Nombre de simulations	600	600

La construction du réseau P2P est réalisée suivant des conditions établies. Les arbres sont constitués en priorité par les groupes qui se trouvent dans le même PID. De même, les groupes sont constitués en priorité par les pairs qui se trouvent dans le même PID. Nous réduisons ainsi le trafic inter-PID en permettant à l'application P2P d'exploiter les informations de localisation fournies par la carte du réseau. Dans les simulations, après la génération d'un nombre aléatoire de clients durant chaque itération, les groupes et les arbres sont formés suivant un ordre aléatoire pour simuler l'arrivée aléatoire de nouveaux utilisateurs dans le réseau. Les connexions entre les arbres sont ensuite décidées par l'application P2P qui exploite les services d'informations ou de sélection offerts par l'opérateur réseau.

4.4.2 Résultats des simulations

Les performances de différentes techniques utilisées pour optimiser les connexions entre les arbres du réseau P2P sont comparées. Ces techniques se basent sur des métriques spécifiques et font intervenir un processus de sélection au niveau applicatif ou bien au niveau réseau.

1. La technique de sélection aléatoire
2. La technique de sélection exploitant des informations de distance
3. La technique de sélection exploitant les coûts de transport
4. La technique de sélection par le réseau, qui exploite les coûts de transport et les met à jour en temps réel.

Les résultats des simulations sont ensuite analysés pour évaluer les performances de transport des flux de données, celles du réseau et celles des services applicatifs comme nous allons le voir.

Taux d'utilisation des liens Les taux d'utilisations des liens du réseau est un indicateur d'une bonne distribution de charge. Il s'agit d'un aspect important à prendre en considération pour exploiter pleinement et de manière efficace l'ensemble des ressources disponibles dans le réseau. Nous présentons les taux d'utilisation des liens au cours une simulation unique dans la figure 4.13 (car il n'est pas possible de faire une représentation de ces valeurs par rapport à l'ensemble des simulations au regard de certains paramètres aléatoires qui varient). Nous remarquons dans cette figure quelques pics d'utilisation qui correspondent à des liens sur-utilisés. L'approche active de collaboration est par ailleurs la seule qui arrive à respecter la limite d'utilisation des liens que nous avons fixé à 80%. L'intérêt qui peut sembler léger dans ce cas particulier est non négligeable. En effet, un lien très utilisé dans le réseau est généralement un lien qui appartient à plusieurs chemins liant différents PIDs du réseau. Si un tel lien est congestionné, toutes les entités applicatives dans tous les PIDs qui exploitent des chemins réseaux incluant ce lien seront affectées. Pour ces raisons, nous partons dans notre approche du constat que chaque lien est important mais aussi que ceux qui sont les plus utilisés sont généralement les

plus importants. En arrivant à limiter l'utilisation des liens et à les contenir sous certains de sécurité, nous assurons une meilleure résilience dans le réseau aux bénéficiaires des applications et de leurs utilisateurs.

Distribution des taux d'utilisation des liens Nous avons comparé la distribution des taux d'utilisation des liens pour les différentes approches par rapport à l'ensemble des simulations que nous avons réalisées avec des profondeurs d'arbres différentes. Ces valeurs fournissent un indicateur de la bonne répartition de la charge dans le réseau. Les figures 4.17, 4.18 et 4.19 illustrent la répartition des taux d'utilisation des liens respectivement pour les applications P2P paramétrées des profondeurs d'arbres de taille 5, 10 et 15. Nous constatons à travers ces résultats que l'utilisation des coûts de transport permet de limiter l'utilisation intensive de certains liens. Ainsi la charge est positionnée sur des liens plus faiblement utilisés. L'écart est notable entre les métriques coût de transport et distance où en moyenne pour chaque simulation 5 liens en plus sont dans la plage d'utilisation entre 60% et 70%. Pour la technique de sélection aléatoire, la différence est considérable. Nous constatons aussi que l'écart entre les différentes techniques a tendance à se réduire quand la taille des groupes d'arbres augmentent. En effet, une taille d'arbre plus importante résulte en un nombre plus important de clients connectés entre eux. Le trafic entre les PIDs est réduit et la charge sur les liens inter-PID du réseau est contenue. En conséquence, les différences entre l'utilisation de telle ou telle métrique deviennent moins prononcées quand la charge de trafic est plus petite. Cependant, le gain réalisé reste intéressant pour différentes raisons. D'un côté, une meilleure utilisation des ressources du réseau, est assurée ce qui permet de supporter une charge supérieure par rapport et d'un autre côté, la qualité de service dans le réseau se trouve améliorée.

Variance des taux d'utilisation des liens du réseau La variance de l'utilisation de liens réseaux est un autre indicateur pour analyser les conséquences de l'utilisation d'une métrique ou technique particulière sur les performances du réseau. En caractérisant la dispersion des taux d'utilisation des liens, la variance permet de remonter les plages supérieures d'utilisation des liens. Nous présentons la variation de la variance des taux d'utilisation des liens pour un ensemble de simulations dans les figures 4.20, 4.21 et 4.22 respectivement pour des tailles d'arbres de 5, 10 et 15 groupes. Le premier sous-ensemble des simulations porte sur des taux initial d'utilisation d'environ 20% et le sous-ensemble restant porte sur un taux d'environ 40%. Ces figures montrent que les variations les plus faibles sont rendues en utilisant la métrique coût de transport. La sélection faite par le réseau permet d'obtenir des résultats légèrement meilleures mais l'utilisation de la métrique coût de transport qui est basée sur les informations de bande passante disponible des liens assure une variation plus faible des taux d'utilisation par rapport à la métrique distance. Cette variance plus faible indique une faible dispersion des taux d'utilisation. Pour des moyennes sensiblement équivalentes, elle assure des pics d'utilisation moins présents que les autres technique ce qui résulte en une meilleure répartition de charge au sein du réseau.

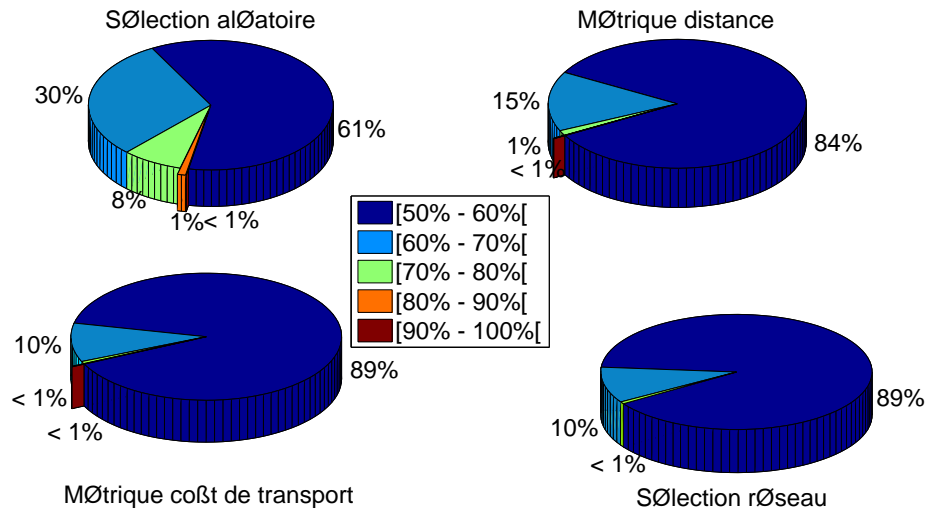


FIGURE 4.17 – La distribution des taux d’utilisation des liens du réseau des différentes techniques de sélection pour une taille d’arbre de 5

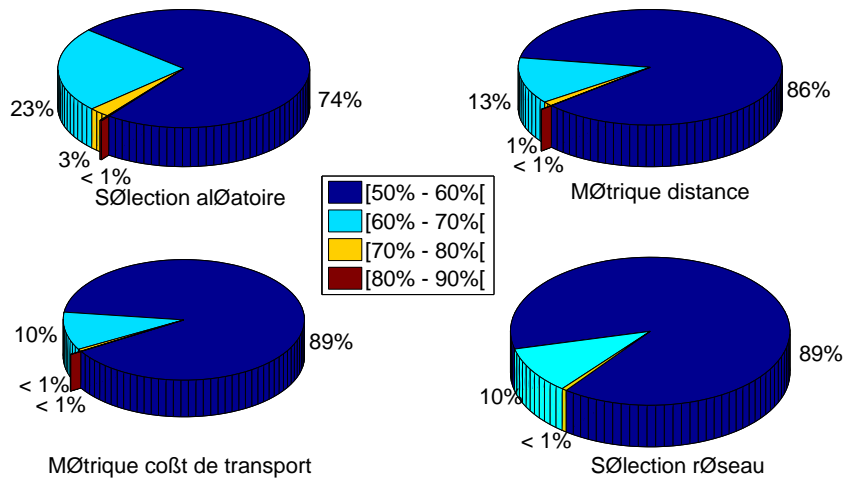


FIGURE 4.18 – La distribution des taux d’utilisation des liens du réseau des différentes techniques de sélection pour une taille d’arbre de 10

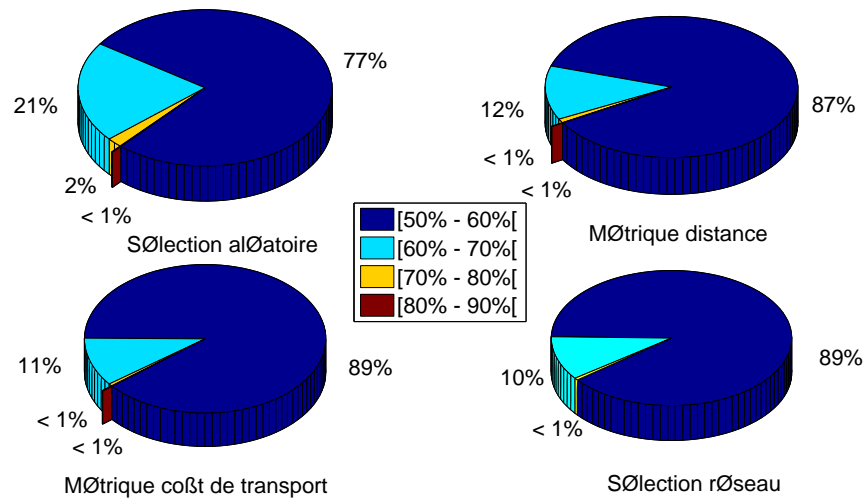


FIGURE 4.19 – La distribution des taux d'utilisation des liens du réseau des différentes techniques de sélection pour une taille d'arbre de 15

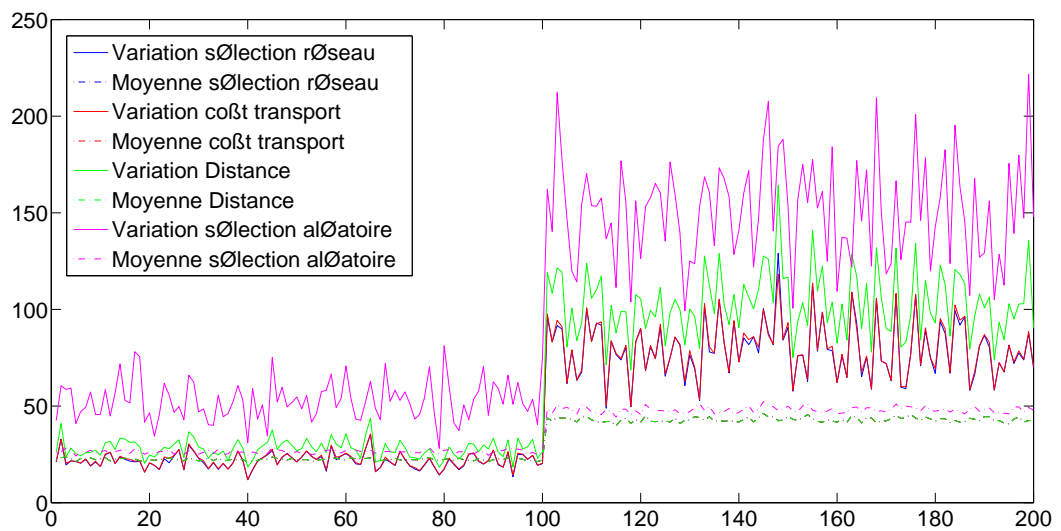


FIGURE 4.20 – Illustration de la moyenne et de la variance des taux d'utilisation des liens pour une taille d'arbre de 5

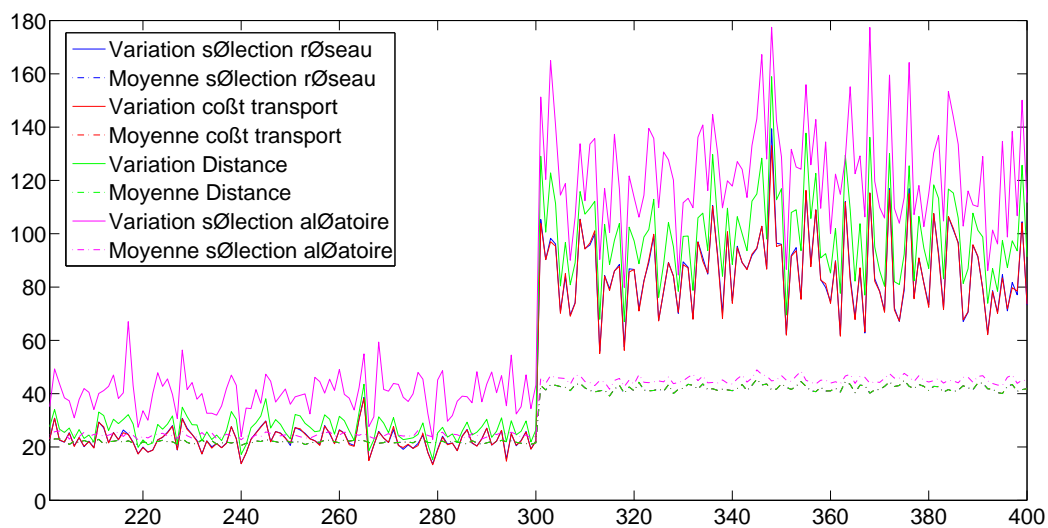


FIGURE 4.21 – Illustration de la moyenne et de la variance des taux d'utilisation des liens pour une taille d'arbre de 10

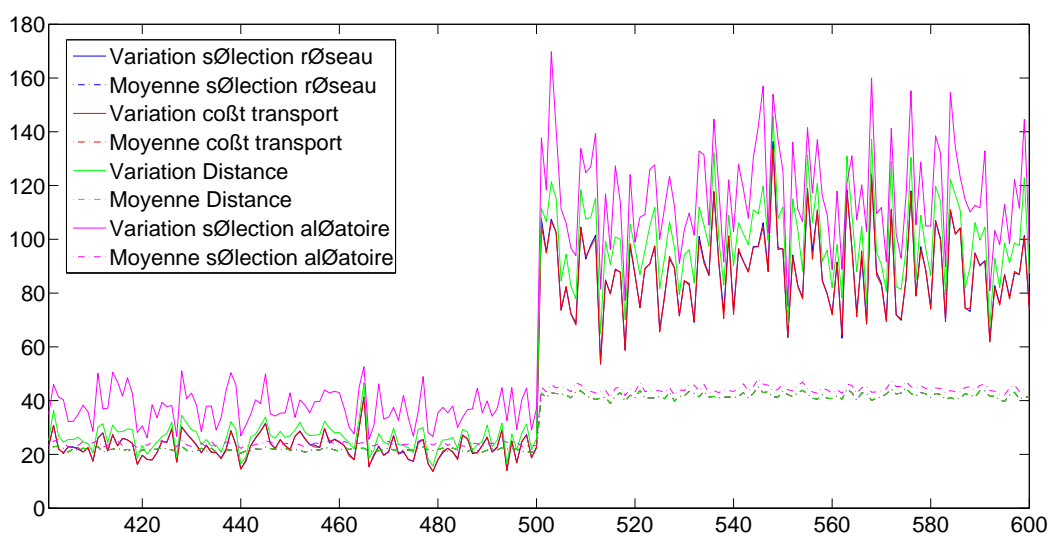


FIGURE 4.22 – Illustration de la moyenne et de la variance des taux d'utilisation des liens pour une taille d'arbre de 15

Gain en utilisation moyenne du réseau Nous présentons dans la figure 4.23 les gains moyens sur l'ensemble des simulations réalisés par chaque technique de sélection par rapport à la sélection aléatoire pour les différentes tailles d'arbres que nous avons modélisées. Notre première constatation porte sur les valeurs de gains réalisés qui oscillent entre 8% et 15%. Pour un trafic, en l'occurrence ici celui de l'application P2P, qui représente entre 4% et 10% du trafic total dans le réseau, ces valeurs sont non négligeables. Associées à d'autres types de service comme les CDNs, les gains cumulés peuvent contribuer à atteindre une meilleure efficacité du réseau qui se traduit en une meilleure qualité pour les services applicatifs.

Si nous comparons par ailleurs les techniques entre elles, nous constatons que l'utilisation de la métrique coût de transport permet de réaliser des gains supérieurs. Contrairement aux résultats que nous avons obtenus dans les simulations que nous avons réalisées avec des services de type client-serveur, la métrique distance n'offre pas toujours le meilleur gain. En effet, dans un réseau P2P, le nombre de sources de données est beaucoup plus important que dans un réseau de distribution de contenus où le nombre de serveurs est limité. Au final, en termes de distance ou nombre de liens utilisés, les deux métriques obtiennent des résultats comparables mais le gain en utilisation penche en faveur de la métrique coût de transport.

La dernière constatation que nous pouvons faire concerne la variation du gain en fonction du nombre de groupes dans chaque arbre. Nous remarquons que le gain tend à augmenter légèrement quand le nombre de groupes augmentent. Ces valeurs s'expliquent par une diminution du trafic sur les liens entre les PIDs au prix d'un retard supplémentaire dans la distribution des données. En effet, les données sont distribuées plus rapidement dans un arbre contenant 5×9 clients que dans un groupes contenant 15×9 . Ces valeurs montrent qu'augmenter la taille des groupes ne permettent pas de réaliser un gain significatif. Mais l'approche peut être intéressante quand le réseau est soumis à une forte charge de trafic.

Latence moyenne réduite La latence moyenne calculée dans le réseau montre que l'utilisation de la métrique coût de transport assure une meilleure fluidité du trafic qui se traduit en un meilleur temps de réponse du réseau. La bonne gestion de l'ensemble des ressources disponibles profite ainsi aux utilisateurs aussi (figure 4.24).

4.4.3 Synthèse et conclusions

Les différents résultats que nous avons présenté dans cette section portent surtout sur les performances de certaines techniques de sélection appliquées à notre modèle P2P de distribution de contenu. Pour choisir une source de données appropriée pour chaque arbre du réseau P2P, nous avons comparé les métriques coût de transport et distance. Ces techniques s'insèrent dans le modèle de collaboration que nous étudions. L'échange d'informations entre le réseau et l'entité de gestion du réseau P2P a permis de réaliser des gains intéressants quant à l'utilisation du réseau, d'améliorer la répartition de la

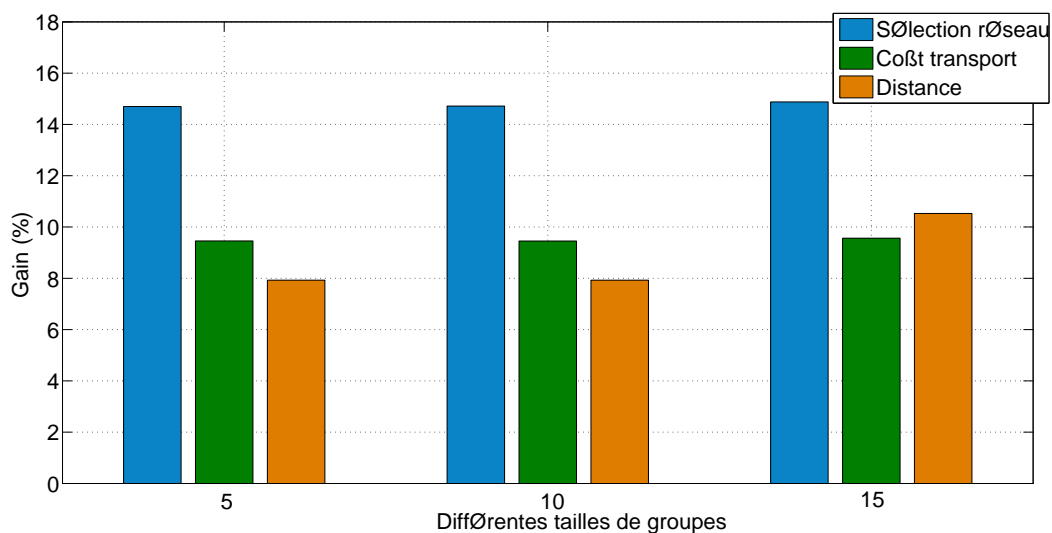


FIGURE 4.23 – Illustration des gains réalisés par rapport à la technique de sélection aléatoire pour différentes tailles d'arbres.

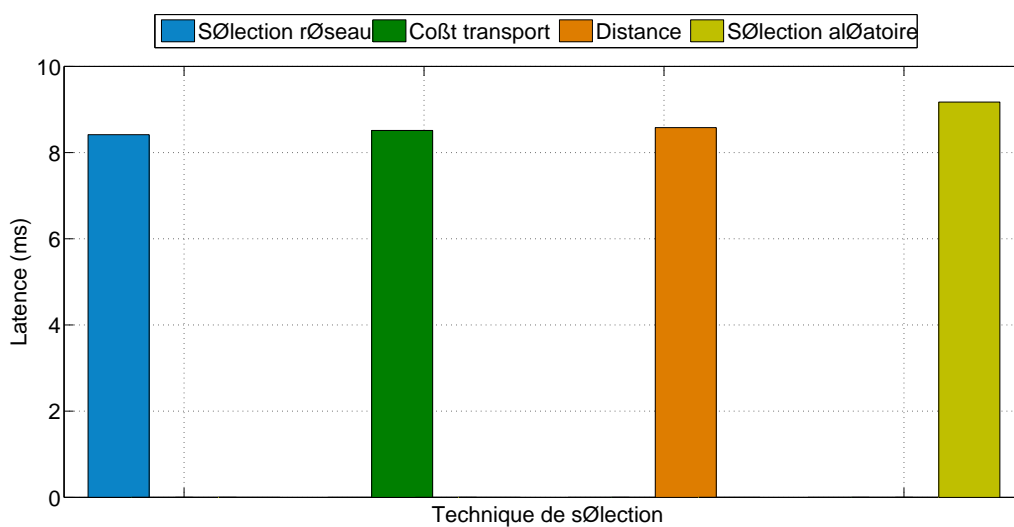


FIGURE 4.24 – Illustration de la latence moyenne dans le réseau pour les différentes techniques

4.4. Évaluation de l'application de l'approche de collaboration à un système P2P

charge du trafic P2P sur l'ensemble du réseau et d'assurer ainsi une meilleure réponse du réseau aux sollicitations des clients P2P.

4.5 Évaluation de l'approche collaborative sur un prototype d'un système de distribution de contenus

Nous présenterons dans cette section les travaux que nous entrepris pour développer un prototype de test sur des systèmes réels. Nous commencerons par présenter le prototype, le contexte et l'environnement des tests menés puis nous détaillerons la procédure des tests. Nous analyserons ensuite les résultats que nous avons obtenus afin de produire une synthèse de l'expérience.

4.5.1 Présentation du prototype de test

Nous avons choisi de développer un prototype d'un système de distribution de contenus en intégrant les fonctionnalités de base. Ces systèmes sont largement utilisés aujourd'hui pour améliorer l'acheminement des contenus vers les utilisateurs finaux et constituent un cas d'application très intéressant pour nos études.

4.5.1.1 Le prototype CDN

Nous nous sommes basés sur architecture simple pour définir le fonctionnement du système. Trois fonctions principales ont été implémentées :

Contrôleur CDN : Cet élément fonctionnel réceptionne les requêtes *HTTP Get* des clients. Il choisit un serveur de distribution suivant une ou plusieurs politiques de décision prédéfinies et renvoie au client une redirection vers le serveur retenu.

Serveur cache CDN : Cet élément est un serveur de distribution qui dispose des contenus susceptibles d'intéresser les utilisateurs. Il reçoit les requêtes des clients et les traite. Dans notre implémentation, cet élément est capable de servir tout type de contenu. Il interagit en *HTTP* avec le client(un navigateur web par exemple). Les contenus vidéo peuvent être transmis à un débit égal à celui de la vidéo pour émuler un service de "*live streaming*".

Client HTTP : Bien que les navigateurs soient suffisants pour interagir avec les autres fonctionnalités du système, nous avons développé notre propre client *HTTP* afin de pouvoir analyser plus finement les détails du flux applicatif. Le client est capable de transmettre des requêtes et traiter des réponses en *HTTP*.

L'ensemble des échanges entre ces entités aboutissant à l'acheminement du contenu au client final est illustré dans la figure 4.25.

4.5.1.2 Le prototype CINA

Le prototype CDN que nous venons de présenter applique des politiques prédéfinies pour le choix d'un cache de distribution à chaque client. Ces politiques s'appuient sur des informations fournies par le réseau d'où le besoin de s'appuyer sur l'architecture CINA pour permettre l'échange d'informations entre le prototype CDN qui fait office de service applicatif et le serveur CINA qui synthétise les informations réseau. Les fonctionnalités implémentées comprennent :

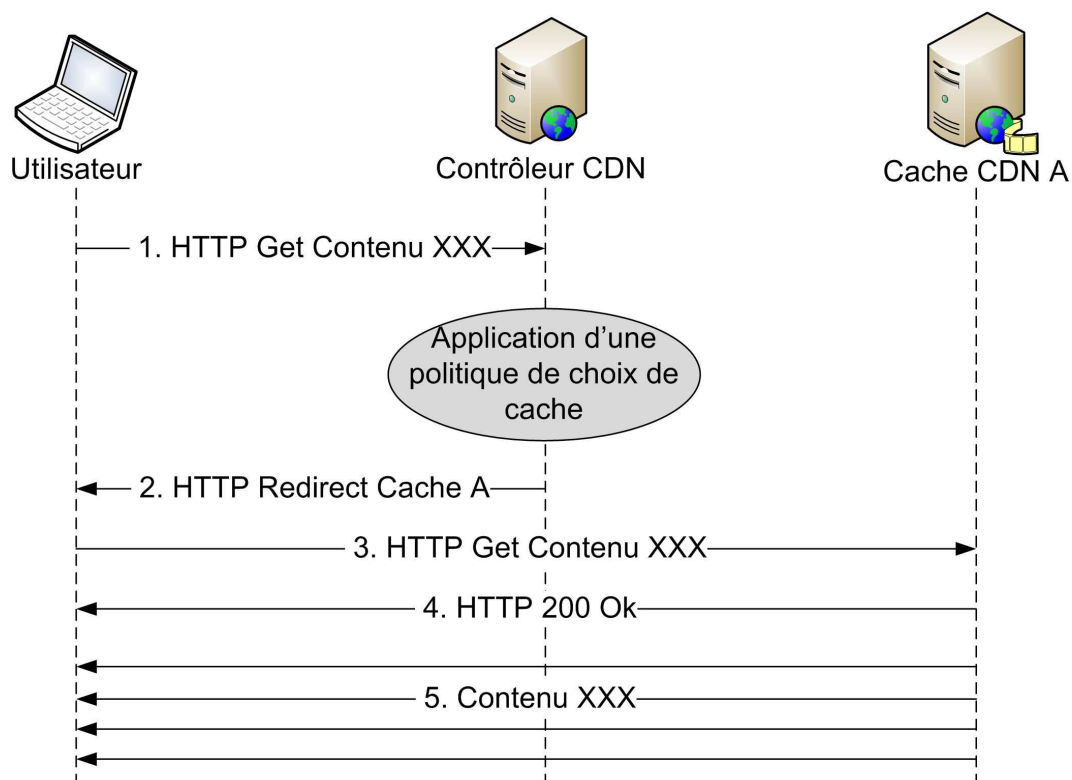


FIGURE 4.25 – Illustration des échanges entre les entités fonctionnelles développées pour le prototype CDN

Serveur CINA : Ce serveur échange avec le service applicatif les informations nécessaires pour l'optimisation du trafic dans le réseau. Le protocole de communication est le protocole ALTO étendu avec les fonctionnalités et informations spécifiques à CINA comme la matrice de proportions. Le protocole d'encapsulation est le protocole HTTP. Dans notre modèle de test, le serveur CINA reçoit une requête du prototype CDN au début de l'expérience pour récupérer les informations pertinentes au réseau.

Client CINA : le client CINA est embarqué dans le système CDN. Il interagit avec le serveur CINA pour transmettre et recevoir des informations et les acheminer à l'entité de décision du Contrôleur CDN.

Les interactions entre l'ensemble des entités fonctionnelles du prototypes sont décrites dans la figure 4.26. L'exécution des différentes phases du système sont indépendantes. En effet, les interactions entre le service CDN et le réseau peuvent être programmées périodiquement dans le temps. Dans notre modèle de test, l'échange d'informations a lieu juste avant de commencer l'exécution du service applicatif (c'est-à-dire l'acheminement des données aux clients qui demandent un contenu). Pendant cette phase, les informations qui seront récupérées seront utilisées par l'algorithme de sélection pour

choisir un serveur cache.

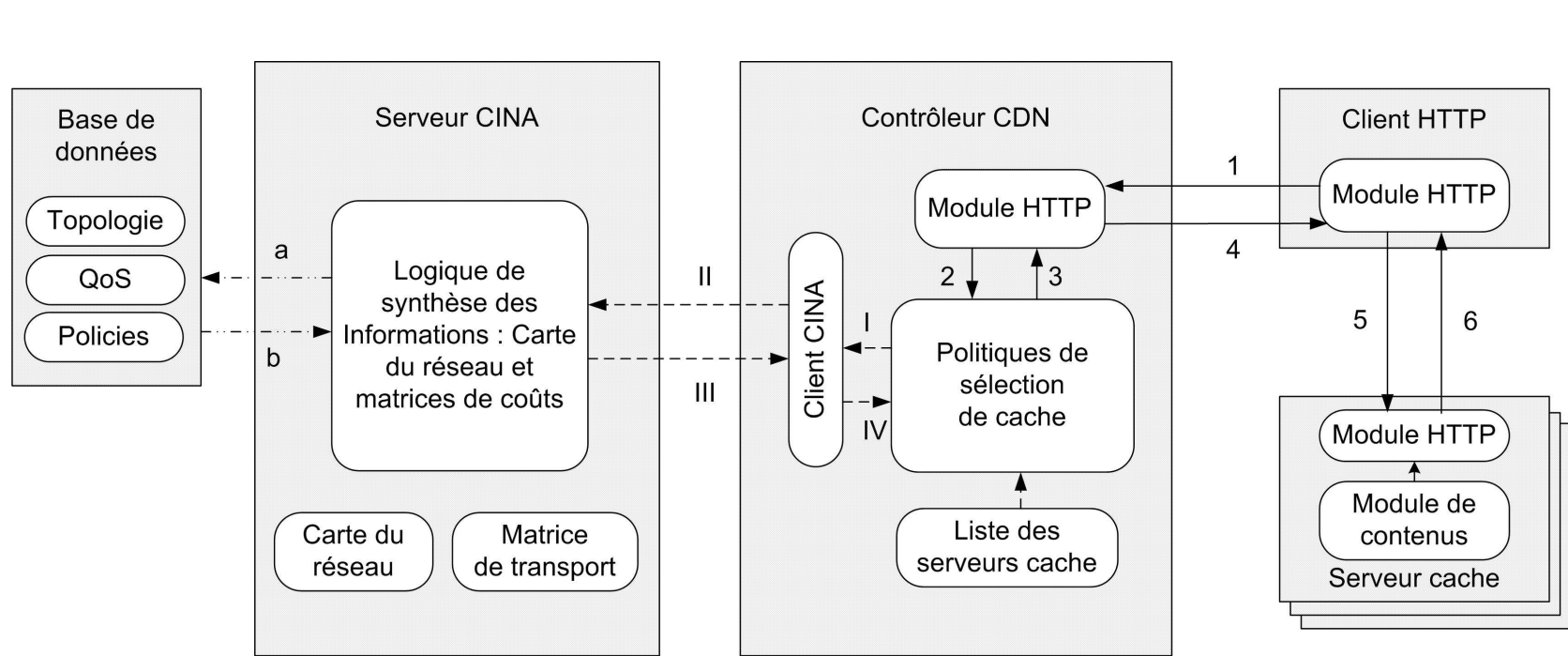
Les étapes [1 – 6] illustrent donc la phase d'exécution du service CDN :

- Etape 1 : Le client demande un contenu au contrôleur CDN via une requête HTTP
- Etape 2 : Le module HTTP du contrôleur CDN demande au module de sélection de choisir un serveur cache pour servir le client
- Etape 3 : Un serveur retenu et son adresse IP est transmise au module HTTP
- Etape 4 : Une redirection HTTP est transmise au client en réponse à sa requête
- Etape 5 : Le client transmet une nouvelle requête au serveur cache
- Etape 6 : Le serveur transmet le contenu demandé au client

Les étapes [I – IV] illustrent l'échange d'informations entre le client et le serveur CINA :

- Etape I : Le module de sélection du contrôleur CDN fait appel au client CINA pour commencer un échange d'informations avec le réseau
- Etape II : Le client CINA prépare et transmet les requêtes au serveur
- Etape III : Le serveur CINA traite les requêtes et envoie les informations aux client.
- Etape IV : Le client traite la réponse du serveur CINA et transmet le contenu au module de sélection qui va les exploiter pour choisir un serveur approprié à chaque client.

Chaque étape indique qu'il y a eu une ou plusieurs interactions entre les entités. Nous ne décrivons pas en détail dans la figure 4.26 l'ensemble des échanges qui ont lieu dans leurs phases respectives mais nous indiquons qu'il y a un acheminement logique d'interactions entre les entités. A titre d'exemple, l'algorithme de sélection peut demander la carte du réseau pour situer un nouveau cache serveur. Il utilise le client CINA pour transmettre la demande au serveur CINA. Le client CINA reçoit les informations en réponse et les fournit au module de décision. Il peut demander par la suite la mise à jour d'une matrice de distance et commencer ainsi un nouvel échange.



4.5. Évaluation de l'approche collaborative sur un prototype d'un système de distribution de contenus

FIGURE 4.26 – Illustration des étapes de fonctionnement du prototype à travers les interactions entre les différentes entités fonctionnelles

4.5.1.3 Le module de monitoring réseau

L'ensemble des informations qui sont rendues disponibles au serveur CINA sont récupérées par un module de monitoring dédié. Ce module dispose des droits accès aux différents équipements réseaux pour réaliser les tâches suivantes :

Construction de la carte du réseau Déterminer l'ensemble des réseaux qui sont desservis par chaque routeur pour construire une carte du réseau où chaque adresse réseau appartient à un PID. Ces informations sont ensuite stockées en base pour permettre de déterminer le chemin entre deux réseaux.

Calcul de la métrique distance Déterminer la distance en termes de nombre de noeuds réseaux entre les PIDs à travers le plan de routage.

Calcul de la métrique latence Déterminer la latence du chemin aller-retour entre tous les PIDs. Cette métrique est calculée en opérant des *ping* directement depuis les routeurs de bout-en-bout entre les PIDs.

L'ensemble de ces informations est stockées dans une base de données consultables par le serveur CINA. Ce dernier ne commande pas le module de monitoring dans notre modèle. L'exécution du module est effectuée avant chaque test pour mettre à jour les informations réseaux. Le serveur CINA récupère le contenu de la base et met à jour les différentes cartes et matrices mises à dispositions des services applicatifs.

4.5.2 Le réseau de test

4.5.3 Présentation de la plateforme de test

Nous avons évoqué en début de ce chapitre les difficultés de réalisation de ce genre de tests qui requièrent la maîtrise d'un réseau étendu et représentatif et la disposition d'une population de clients et d'un ensemble de serveurs pour pouvoir évaluer efficacement nos travaux. Nous avons opté pour l'utilisation d'une plateforme d'expérimentation d'Orange qui offre une large étendue en inter-connectant des sites à Lannion, à Rennes et à Paris. Les équipements réseaux sont similaires aux équipements présents dans le réseau de production car utilisés à des fins de tests. La plateforme offre différents types de réseaux d'accès comme l'ADSL, la fibre ou les réseaux mobiles. Certaines mesures de sécurité et certaines contraintes logistiques ne permettaient pas une exploitation parfaite de la plateforme. Il n'était pas possible ainsi de connecter directement des équipements comme les serveurs aux équipements réseau mais il fallait passer par des agrégations de vlans mise à disposition de notre laboratoire. La deuxième contrainte concernait le nombre de clients simultanés que nous pouvions avoir pendant chaque test. Nous avons dû équiper des serveurs physiques de plusieurs cartes réseau. Nous avons exécuté ensuite un certain nombre de clients sur la machine en les répartissant sur l'ensemble des interfaces réseau. Chaque interface représente ainsi un PID et les clients qui l'utilisent appartiennent à ce PID. La plateforme de test utilisée est présentée dans la figure 4.27. L'ensemble des paramètres de la plateforme et du service est détaillé dans la table 4.5.

4.5. *Évaluation de l'approche collaborative sur un prototype d'un système de distribution de contenus*

TABLE 4.5 – Paramètres des tests

Paramètre	Détail	Localisation	Machine physique
PID1	10 clients	Lannion	1
PID2	10 clients	Lannion	1
PID3	10 clients	Lannion	2
PID4	Contrôleur CDN	Lannion	4
PID5	10 clients	Paris	3
PID6	Serveur cache	Lannion	4
PID7	10 clients	Lannion	1
PID8	10 clients	Lannion	2
PID9	Serveur cache	Lannion	5
PID10	Serveur cache	Paris	6

TABLE 4.6 – Configuration des services CDN

PID	CDN-A			CDN-B		
	Clients	Serveurs	Contrôleur	Clients	Serveurs	Contrôleur
1	5	0	0	5	0	0
2	5	0	0	5	0	0
3	5	0	0	5	0	0
4	0	0	1	4	0	1
5	5	0	0	5	0	0
6	0	1	0	0	0	0
7	5	0	0	5	0	0
8	5	0	0	5	0	0
9	0	1	0	0	1	0
10	0	0	0	0	1	0

4.5.3.1 La configuration des services CDN

Nous avons apporté une attention particulière depuis les débuts de nos travaux sur la collaboration entre applications et réseaux sur la gestion de plusieurs services applicatifs simultanément, pour garantir une meilleure application dans un environnement de production et atteindre de meilleures performances. Dans cette optique, nous avons défini dans notre modèle de test deux services CDN indépendants disposant chacun de son contrôleur, ses serveurs cache et ses clients (table 4.6).

4.5.3.2 Les résultats des redirections des clients

Nous avons appliqué quatre politiques de sélection différentes pour la redirection des clients vers les serveurs de chaque CDN :

Redirection circulaire : Le contrôleur sélectionne un serveur de façon circulaire dans la liste de serveurs disponibles à chaque requête client. Cette approche permet

de balancer la charge équitablement sur l'ensemble des cache CDN disponibles dans le réseau sans se soucier des conditions de transport.

Utilisation de la matrice de coût distance .

Utilisation de la matrice de coût latence .

utilisation de la matrice de proportions : Cette politique de sélection applique l'approche de collaboration active pour déterminer l'ensemble des choix optimaux en fonction du nombre et des localisations des clients et des serveurs. En ignorant les conditions réelles de trafic dans la plateforme de test, nous avons opté de partir d'un trafic initial nul dans nos calculs.

Nous résumons le résultats des redirections des clients pour chaque approche pour le CDN-A dans la table 4.7 et pour le CDN-B dans la table 4.7.

TABLE 4.7 – Présentation des résultats des redirections des clients du CDN-A vers les serveurs localisés dans les PIDs 6 et 9

PID Client	Sélection		Métrique		Métrique		Sélection	
	Circulaire		Distance		Latence		Optimale	
	PID6	PID9	PID6	PID9	PID6	PID9	PID6	PID9
1	3	2	0	5	5	0	3	2
2	3	2	0	5	5	0	4	1
3	1	4	0	5	5	0	1	4
5	2	3	0	5	5	0	4	1
7	0	5	5	0	5	0	4	1
8	4	1	5	0	5	0	1	4

TABLE 4.8 – Présentation des résultats des redirections des clients du CDN-B vers les serveurs localisés dans les PIDs 9 et 10

PID Client	Sélection		Métrique		Métrique		Sélection	
	Circulaire		Distance		Latence		Optimale	
	PID9	PID10	PID9	PID10	PID9	PID10	PID9	PID10
1	2	3	5	0	5	0	4	1
2	4	1	5	0	5	0	4	1
3	3	2	5	0	5	0	1	4
5	3	2	0	5	0	5	2	3
7	0	5	5	0	5	0	2	3
8	3	2	5	0	5	0	4	1

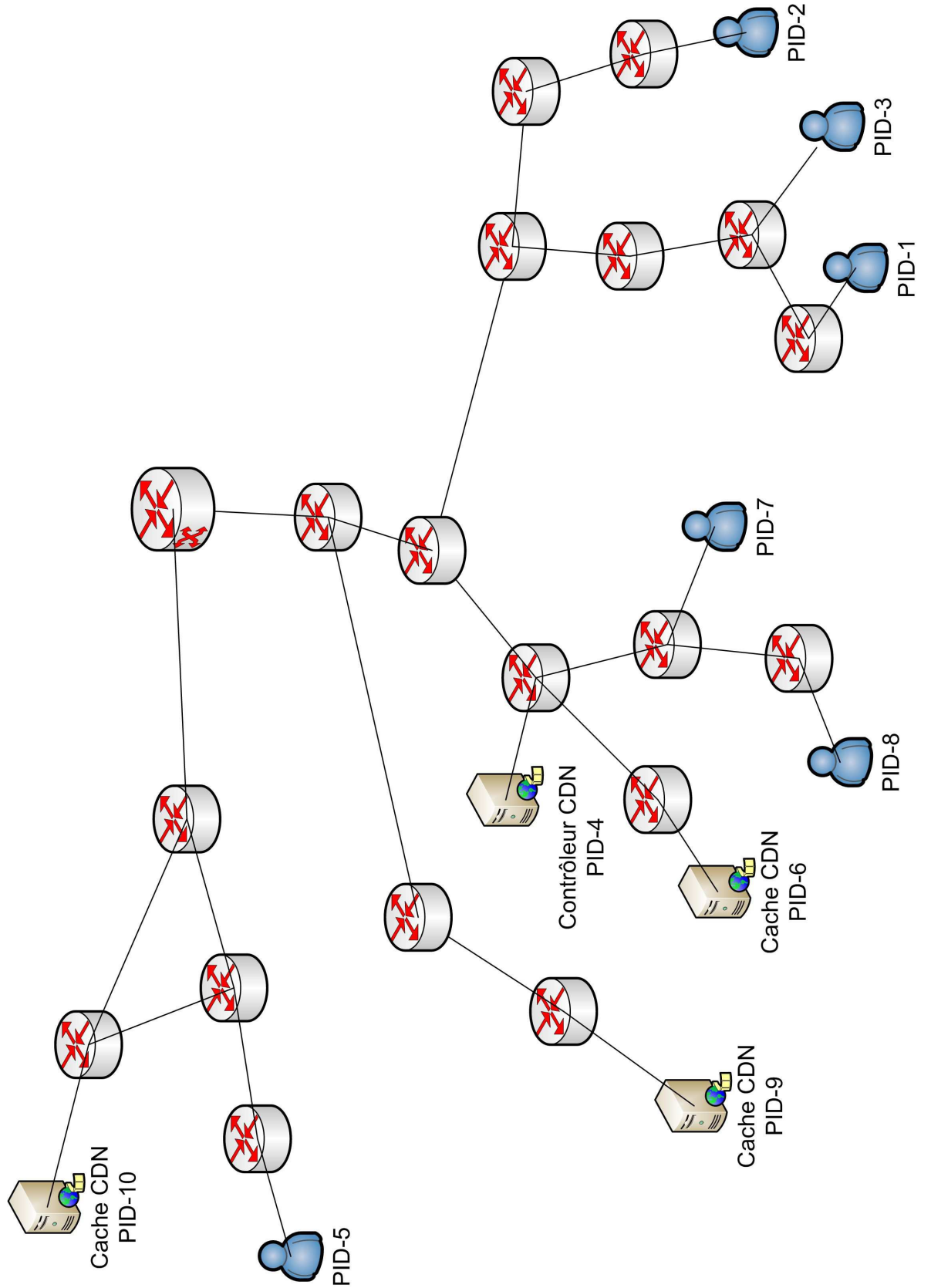


FIGURE 4.27 – Illustration de la plateforme de tests avec les différentes entités du service applicatif

4.5.4 Les résultats des tests

Nous présentons dans cette section les résultats d'exécution du prototype que nous avons développé sur la plateforme de test présentée précédemment. Nous avons procédé à la capture du trafic relatif aux services CDN sur chaque interface réseau pour l'analyser. Nous avons enregistré aussi les informations d'exécution de chaque client applicatif *HTTP* que nous avons lancé pour déterminer les instants de réception des données. Ces données fournissent toutefois des informations redondantes avec celles fournies par la capture du trafic plus bas dans la couche réseau.

4.5.4.1 La mesure des débits applicatifs de chaque PID

Nous avons analysé les sessions TCP des utilisateurs des deux services applicatifs pour mesurer les débits de données dont a bénéficié chaque PID associé à une politique de sélection spécifique indépendamment du service. Le débit applicatif est un indicateur de la capacité du réseau à absorber efficacement les données générées par les deux services CDNs. Ce débit dépend directement du chemin réseau emprunté par le trafic car généralement il est limité par la portion la moins fluide du chemin qui ne peut supporter un débit supérieur. Nous présentons les débits applicatifs des PIDs clients dans les figures 4.28 à 4.33.

Les résultats obtenus montrent que la politique de sélection optimale, qui s'appuie sur l'approche de collaboration active entre les réseaux et les deux services applicatifs, permet d'entretenir souvent des débits légèrement supérieurs à ceux des autres techniques de sélection. Dans le même registre, la durée moyennes des différentes sessions TCP des utilisateurs de chaque PID est plus courte quand la sélection optimale est employée ce qui va de pair avec un débit applicatif légèrement supérieur. Quand aux autres politiques de sélection, les résultats montrent que les techniques ont produit des performances sensiblement équivalentes. Il y a quelques différences qui se remarquent entre les différents PIDs avec un léger avantage pour l'une ou l'autre. Les politiques de sélection basée sur les métriques distance et latence offrent ainsi de meilleures performances dans les PIDs 7 et 8 grâce à une localisation favorable, proche d'un serveur cache. La sélection circulaire de serveurs a permis cependant d'offrir des débits équivalents pour les PIDs 1, 2 et 3 grâce à une répartition du trafic sur les différents serveurs de chaque service prototype CDN. Cette répartition, non optimale et pas toujours garantie, assure une distribution de charge sur plusieurs chemins réseaux contrairement aux métriques distance et latence qui s'emploie à utiliser l'unique chemin qui minimise la métrique. Nous rappelons toutefois que ces résultats encourageants restent insuffisants pour jauger de façon indéniable les performances des différentes approches que nous avons comparé. En effet, les contraintes de l'accès indirect à travers des agrégations de vlans à la plateforme de test peuvent fausser les résultats de performance.

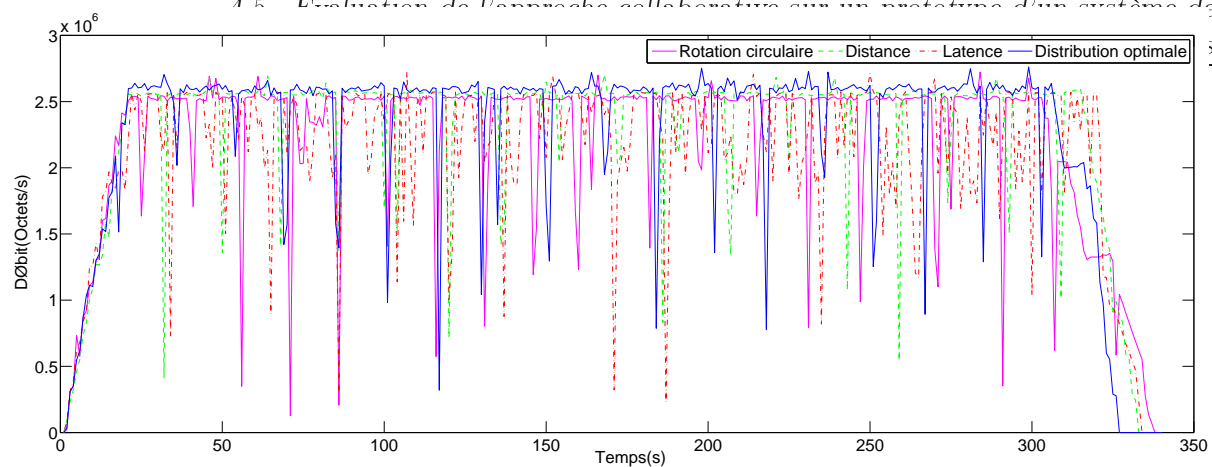


FIGURE 4.28 – Débit du service applicatif du PID 1

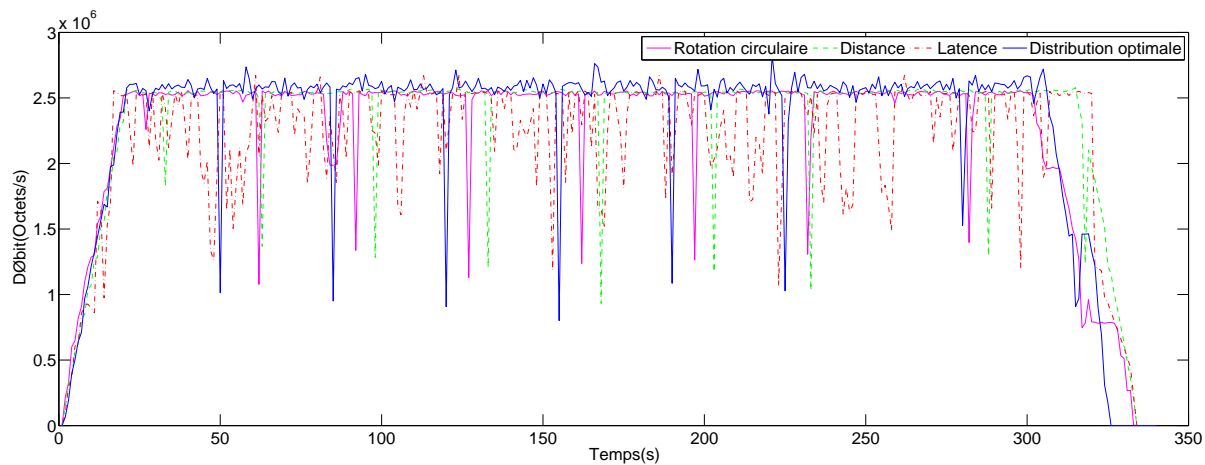


FIGURE 4.29 – Débit du service applicatif du PID 2

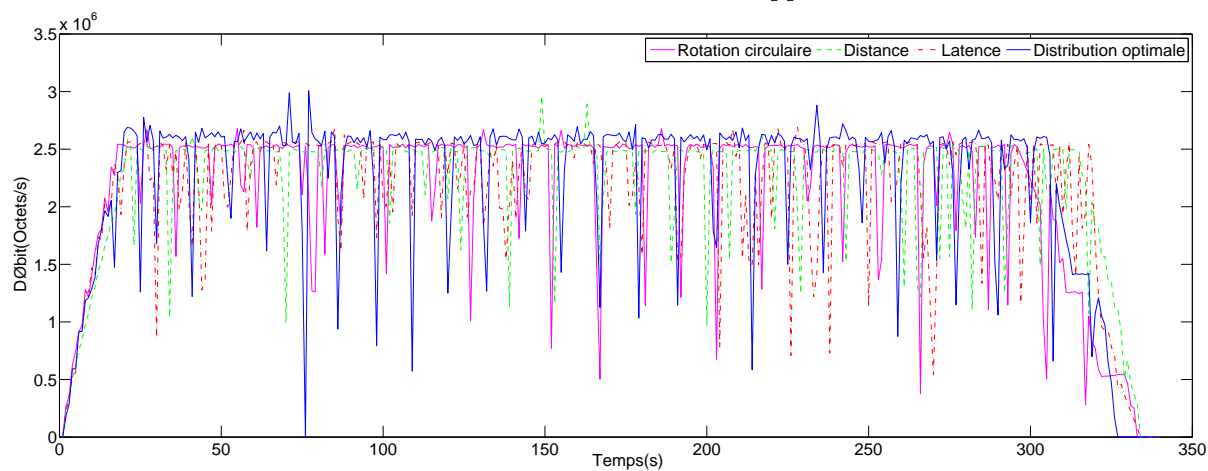


FIGURE 4.30 – Débit du service applicatif du PID 3

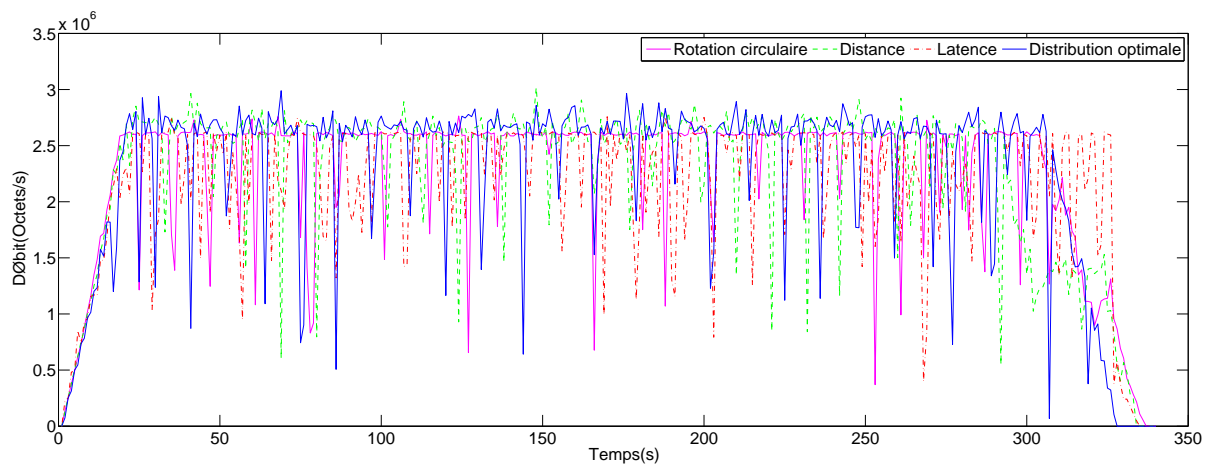


FIGURE 4.31 – Débit du service applicatif du PID 5

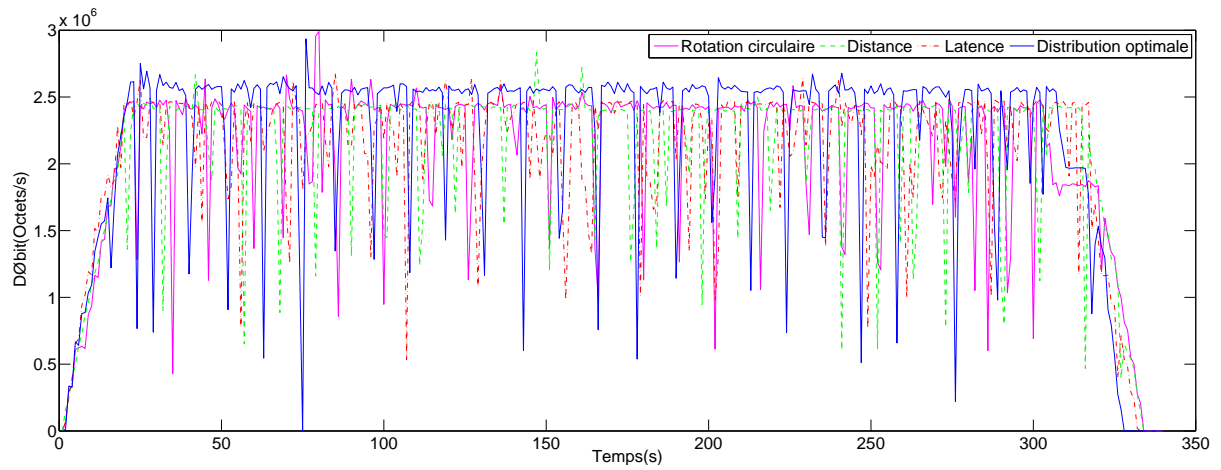


FIGURE 4.32 – Débit du service applicatif du PID 7

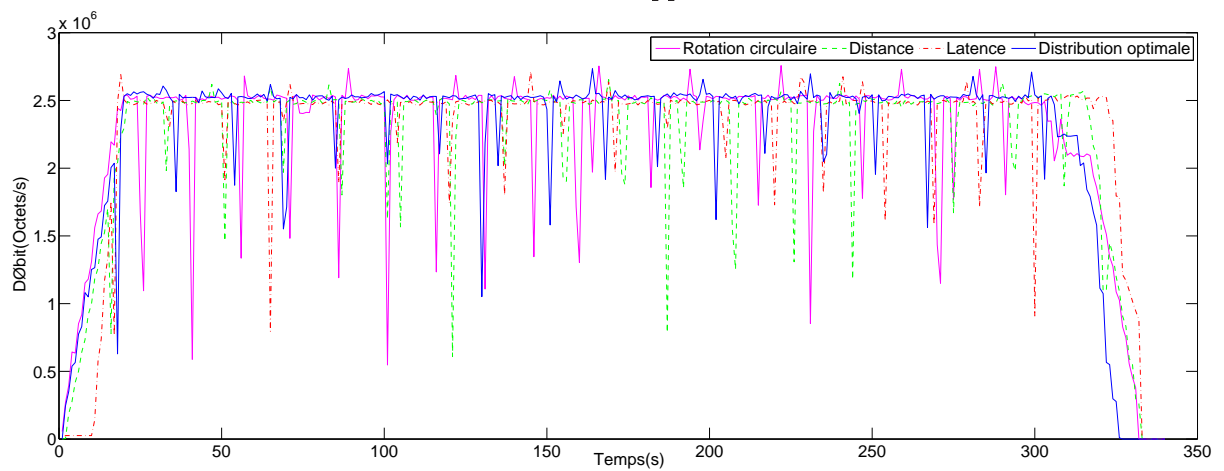


FIGURE 4.33 – Débit du service applicatif du PID 8

4.5.4.2 La mesure d'autres indicateurs de performance

Le RTT moyen par PID Nous avons mesuré aussi les RTT des sessions TCP dans les différents PID à travers les captures de trafic effectuées. Nous voulions mesurer un RTT moyen dans le temps pour chaque PID afin de comparer les différentes politiques de sélection. Nous présentons ces résultats dans les figures 4.34 à 4.39. Il est cependant plus difficile de se prononcer sur l'efficacité de cet indicateur que celui du débit applicatif. En effet, nous nous sommes assurés au préalable que nos accès à la plateforme de test sont capables de supporter la charge de trafic que nous générions avec nos prototypes CDN. Ainsi, le débit applicatif ne se trouve pas handicapé par les chemins d'accès à la plateforme. Nous ne pouvons faire de même pour la latence sur ces chemins. Dans ce contexte, malgré que dans certaines figures, nous pouvons apercevoir une performance légèrement meilleure pour la technique de sélection optimale, une analyse prudente de notre environnement de test nous empêche de tirer des conclusions de ces résultats.

D'autres indicateurs Toujours soumis aux mêmes contraintes, d'autres indicateurs comme le taux de perte de paquets, le taux de paquets désordonnées ou bien le nombre de retransmissions sont autant de mesures que nous sommes capables de présenter mais dont la valeur est altérée par les tronçons communs d'accès à la plateforme de tests. Cette topologie particulière nous empêche de savoir quelle est l'origine de perte d'un paquet par exemple. Nous nous sommes abstenus de présenter ainsi ces figures.

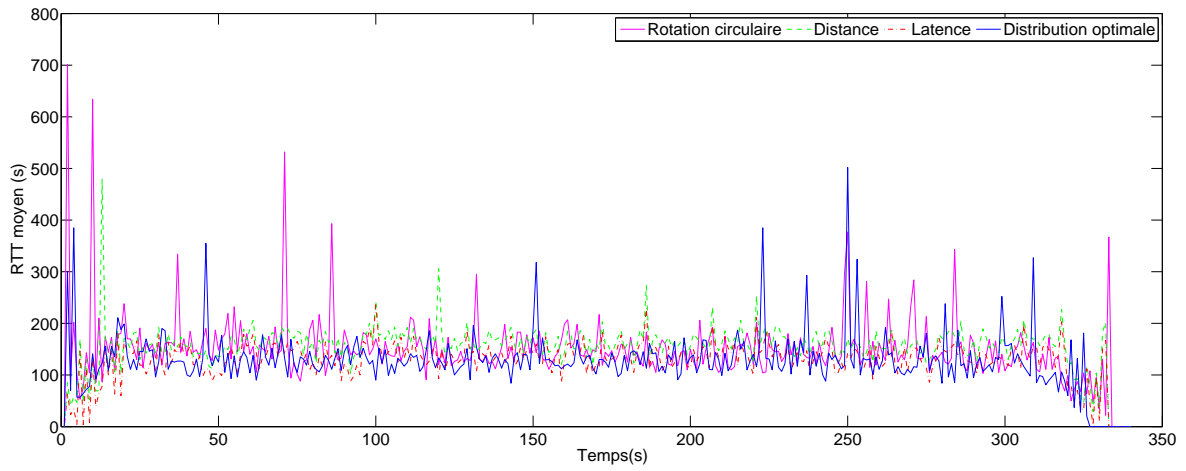


FIGURE 4.34 – RTT moyen pour les sessions des utilisateurs du PID 1

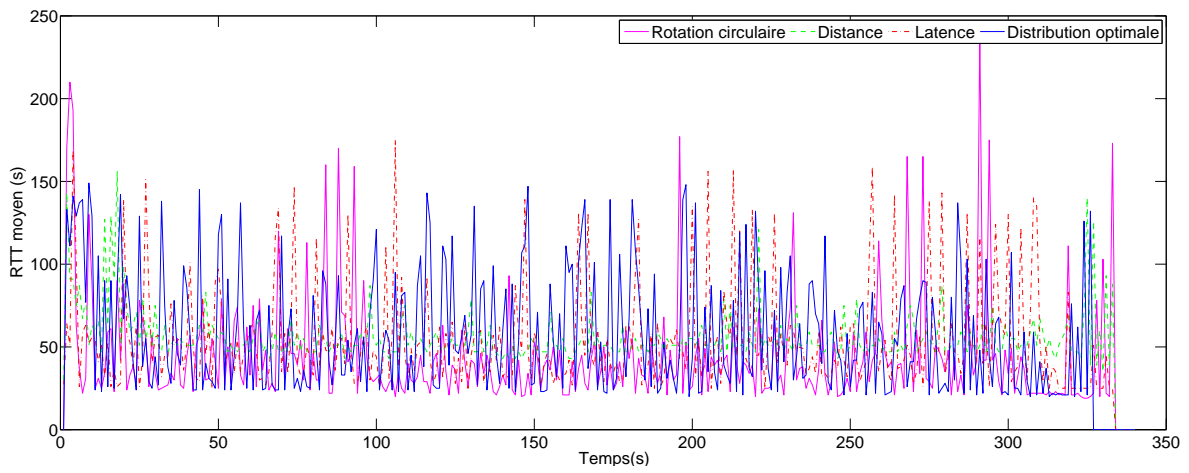


FIGURE 4.35 – RTT moyen pour les sessions des utilisateurs du PID 2

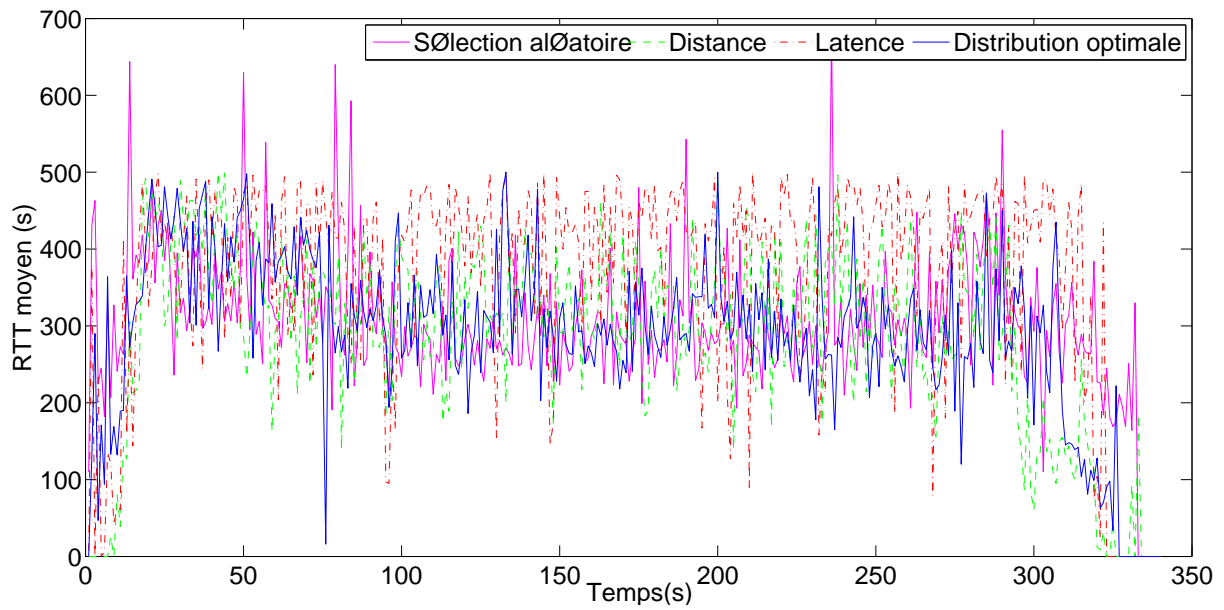


FIGURE 4.36 – RTT moyen pour les sessions des utilisateurs du PID 3

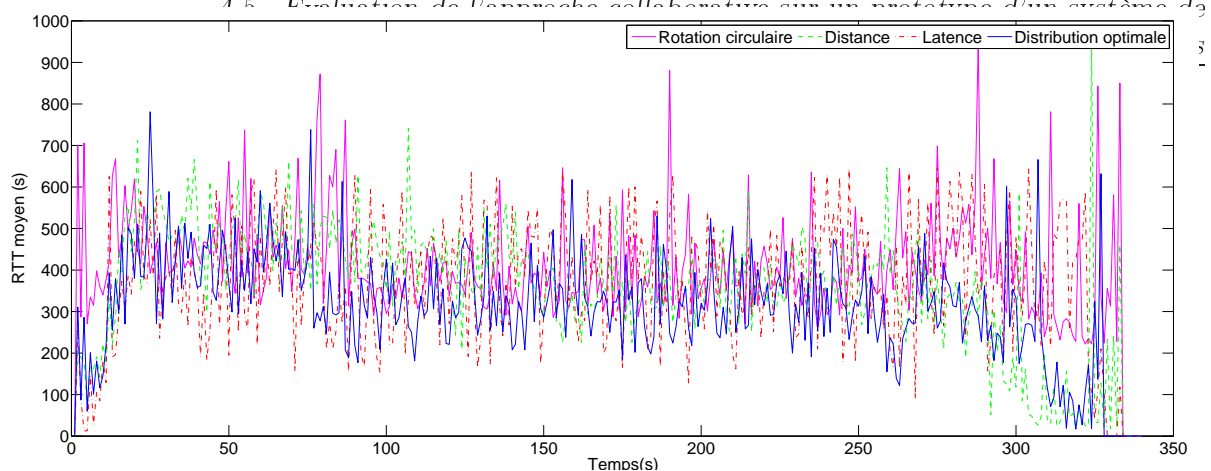


FIGURE 4.37 – RTT moyen pour les sessions des utilisateurs du PID 5

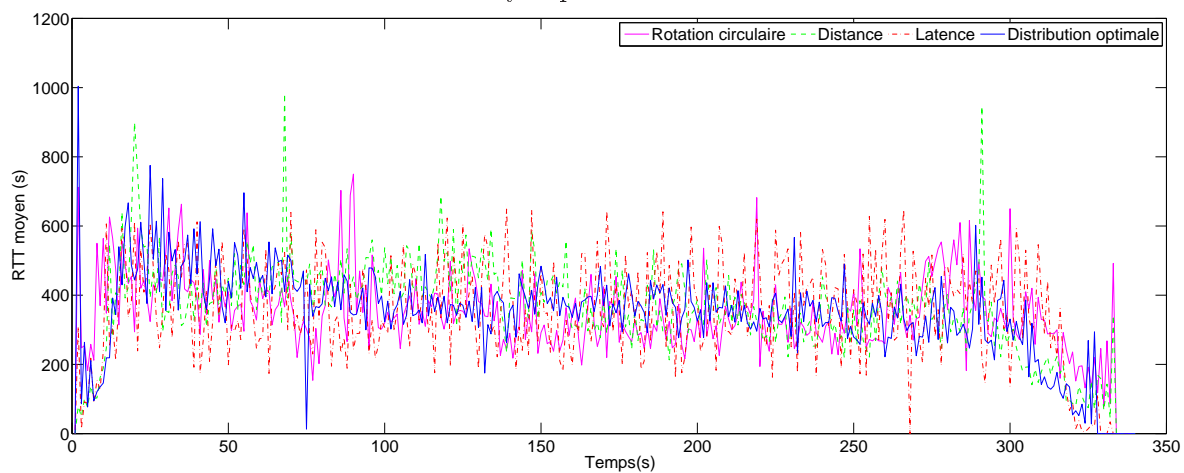


FIGURE 4.38 – RTT moyen pour les sessions des utilisateurs du PID 7

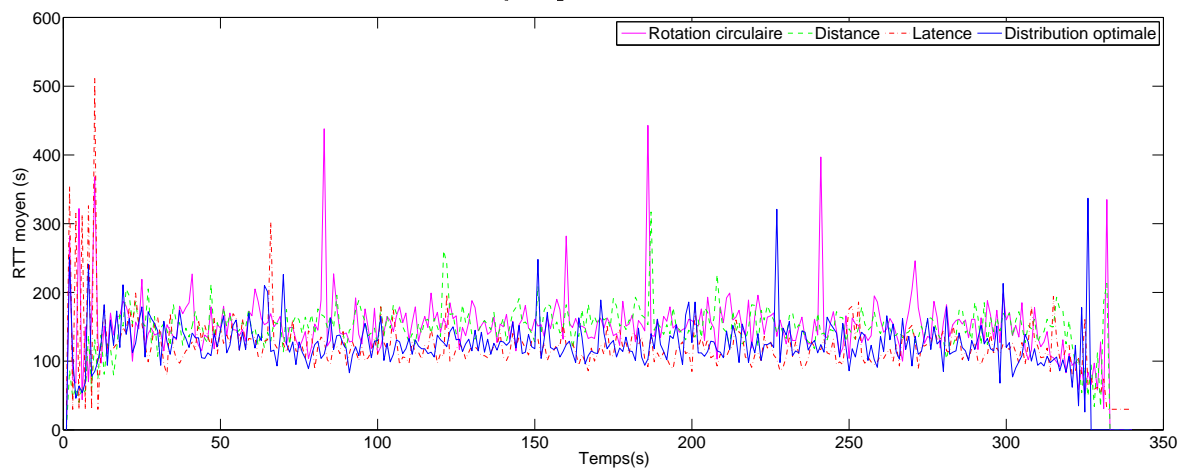


FIGURE 4.39 – RTT moyen pour les sessions des utilisateurs du PID 8

4.5.4.3 Analyse de la répartition de la charge dans le réseau

Nous avons calculé le taux d'utilisation des différents liens utilisés dans la plateforme à travers la connaissance des chemins réseaux empruntés par le trafic. Nous avons supposé un trafic initial nul dans le réseau pour deux raisons. La première raison est la capacité de transfert des liens et équipements de la plateforme qui dépasse souvent $1Gbps$. Nous avons pris cette valeur comme référence dans nos calculs. La seconde raison est la réalisation des différents tests pendant des heures de faible utilisation du réseau pour éviter les perturbations.

Nous présentons donc la répartition des taux d'utilisation des liens, qui est une faible utilisation entre 0 et 8% environ, dans la figure 4.40. Nous pouvons voir que la sélection optimale a permis de contenir l'utilisation des liens en dessous de 4%, ce qui équivaut à un débit effectif de $40Mbps$. Les autres techniques de sélection présentent un certain nombre de liens qui dépassent cette valeur d'utilisation. C'est la technique qui exploite les métriques de latence qui obtient les taux d'utilisation les plus élevés. En effet, à travers les résultats des redirections des clients présentés dans les tables 4.7 pour le CDN-A et 4.7 pour le CDN-B, nous remarquons que le plus grand nombre de clients redirigés ensemble vers le même serveur (55 clients en tout entre les deux CDNs sont redirigés vers le serveur dans le PID 6) est réalisée par cette politique de sélection. Ces choix mènent donc évidemment à une utilisation plus intensive de certains liens, même si dans notre plateforme de test, ces valeurs d'utilisation ne présentent pas de besoins d'actions.

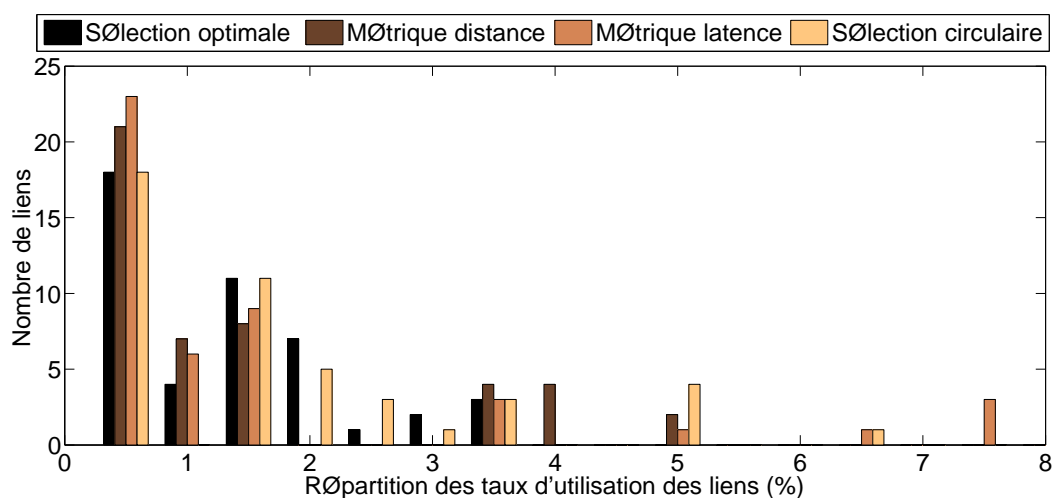


FIGURE 4.40 – Répartition des taux d'utilisation des liens pour les différentes techniques de sélection

4.5.5 Conclusion et synthèse

Les développements et tests que nous avons menés durant nos travaux nous ont permis de valider le fonctionnement de certains systèmes comme l'interface CINA et de mesurer l'importance de la chaîne de bout-en-bout depuis la surveillance du réseau jusqu'à l'utilisation des informations par les services applicatifs. Si certains résultats obtenus se sont avérés difficilement exploitables, il n'en reste néanmoins que les prototypes sont aptes à des tests de plus grande envergure et dans un environnement plus adapté.

Conclusion

Nous avons consacré nos efforts durant cette thèse au développement d'une nouvelle vision sur les relations entre opérateurs réseaux et fournisseurs de services. Nous espérons à travers ces travaux contribuer à une nouvelle hégémonie dans la gestion du transport de données basée sur une collaboration mutuelle et fructueuse. Les approches de collaboration active et passive que nous proposons sont les pierres angulaires de cette hégémonie. Elles se basent sur une interface de collaboration dont l'essence est tirée du monde IT avec pour objectif de respecter trois points fondamentaux :

- Simplicité : pour permettre une adoption facile à très large échelle et sans mobilisation de lourds investissements ni efforts. L'interface CINA à ce titre, extension du protocole ALTO lui même protocole JSON encapsulé dans HTTP, est un exemple d'ingéniosité dans la réalisation simple et évolutive de framework de communication.
- Efficacité : la collaboration n'a de sens que si elle apporte une valeur ajoutée et de manière efficace. Cette efficacité se mesure par la capacité de nos techniques à répondre à des besoins avérés en QoE pour les services et en meilleure gestion des ressources pour les réseaux à travers une répartition plus fine du trafic dans le réseau. Pour atteindre ces objectifs, nous avons porté énormément d'attention dans nos travaux à la prise en compte à la fois des besoins mais aussi des contraintes de chaque acteur afin d'apporter une réponse efficace à la problématique de chaque acteur.
- Adéquation : Le cadre d'exploitation de toute technique est un paramètre fondamental qui doit être tenu en considération dans sa conception. Pour cette raison, notre approche a été adaptée aux exigences d'acteurs différents afin de permettre et faciliter son exploitation. Nous pouvons que la répartition du trafic dans notre modèle est réalisée au niveau applicatif par les applications qui en sont la source. En effet, cette approche est adaptée au contexte actuel où les services sont proposés par des acteurs internet dont le trafic est étranger aux opérateurs réseaux. En conséquence, une gestion au niveau applicatif de ce trafic est plus adaptée qu'une intervention dans le réseau pour des différentes raisons techniques, légales et financières. Cependant, les décisions sont prises soient au niveau applicatif soit au niveau réseau suivant l'approche de collaboration adoptée.

Ces points fondamentaux ont été pris en considération depuis les travaux d'élaboration de l'architecture fonctionnelle et la spécification de l'interface. Les efforts qui ont

été investis dans ces tâches nous permettent d'aujourd'hui de proposer des techniques qui avancent des avantages indéniables en termes de quantité de ressources nécessaires aux déploiements et de facilité d'exploitation outre les performances atteintes.

Les deux approches de collaboration proposées offrent par ailleurs plus de flexibilité dans les cadres de collaboration bilatéraux entre opérateurs réseaux et fournisseurs de services. Les contraintes et les besoins de chacun permettent de définir le niveau de collaboration à retenir pour achever une meilleure gestion du trafic injecté dans le réseau et en conséquence une meilleure QoE pour les utilisateurs. Si les performances atteintes constituent l'intérêt majeur de l'emploi de notre approche de collaboration, il n'en reste pas moins que permettre un échange d'informations entre des acteurs réseaux et services de façon simple et sécurisante est une approche très prometteuse à l'heure de la restructuration de la chaîne de valeur de distribution des contenus. Dans cette perspective, l'interface CINA a été pensée pour permettre l'exploitation de certains services réseaux par les applications en vue de pousser encore plus loin les améliorations des performances aussi bien côté réseaux que côté applications. Ainsi, un service de transmission multicast qui substituerait à un mode de distribution en P2P est capable d'alléger la charge de trafic dans le réseau et d'améliorer la qualité de service dont bénéficie l'application.

Ces performances que nous évoquons ont été présentées et analysées pour évaluer nos propositions à travers des simulations analytiques et des expérimentation réelles. Les résultats obtenus sont très prometteurs et annoncent des gains réalisés non négligeables de l'ordre de 40 à 50 % pour les opérateurs réseaux et une nette amélioration des performances de transport pour les services applicatifs. Les expériences effectuées dans un réseau physique ont permis de consolider les résultats obtenus.

Les différents travaux que nous avons entrepris durant la thèse ont fait l'objet de plusieurs publications dont trois internationales. Dans [MES⁺11b], nous présentons le développement de l'interface de collaboration CINA en cours de spécification dans le projet ENVISION. Ensuite, nous présentons notre architecture fonctionnelle dans [Eil12]. Nous détaillons les différents modules fonctionnels et nous introduisons le modèle d'échange mutuel d'informations. Puis nous développons le modèle collaboration entre réseaux et applications qui permet de profiter des services réseaux mis à disposition. Dans [EMT12], nous avons présenté notre modèle de collaboration passive entre réseaux et applications. Les simulations que nous avons effectuées ont porté sur la comparaison de différentes métriques d'intérêt pour les opérateurs réseaux. Notre modèle de collaboration active est présenté dans [EMT13]. L'approche de collaboration active est comparée à l'approche de collaboration passive afin d'évaluer les performances de cette proposition par rapport aux contraintes qu'elles introduit.

En plus de ces publications, nous avons contribué aux travaux d'études et de spécifications des interfaces d'inter-connexions entre CDNs dans le groupe de travail CDNi à

IETF [SE12]. Nos études nous ont permis par ailleurs de déposer quatre demandes de brevets à l'INPI dont deux sont en cours d'extension aux Etats-Unis et dans d'autres régions du monde et deux en cours d'instruction.

Perspectives et évolutions L'aboutissement ultime des travaux que nous avons entrepris dans un milieu professionnel au sein d'Orange Labs à Lannion est un déploiement dans un réseau expérimental représentatif des réseaux d'Orange. Aujourd'hui, nous avons pu démontrer à travers un prototype la réelle possibilité de continuer sur cette voie et les résultats que nous avons obtenus traduisent un intérêt indéniable. Ainsi, nous espérons pouvoir porter ces travaux sur le terrain. Nous prospectons la possibilité de les adapter à un nouvel environnement d'utilisation. La surveillance passive du réseau peut être remplacée par des mesures actives effectuées à partir de sondes placées stratégiquement dans les zones névralgiques du réseau. Par ailleurs, notre approche est sujette à divers évolutions. Nous pensons particulièrement à la possibilité d'adapter la technique et la métrique utilisées dans la gestion du trafic à un certain nombre de paramètres comme l'état de charge du réseau, le type de trafic, le niveau de collaboration.

Le plus réjouissant quant aux travaux de cette thèse est sans doute de savoir que cette vision de collaboration et d'ouverture des réseaux est entrain de s'imposer au sein de la communauté Télécom et IT.

Glossaire

AAA Authentication, Authorisation and Accounting. 63

ADSL Asymmetric Digital Subscriber Line. 111

ALTO Application Layer Traffic Optimisation. 19

API Application Programming Interface. 48

AS Autonomous System. 26

BCMP Baskett, Chandy, Muntz, Palacios. 120

CDN Content Delivery Networks. 17

CINA Collaboration interface between Networks and Applications. 85

CUT Capacité Unitaire de Transmission. 111

DC Data Centers. 17

ECN Explicit Congestion Notification. 25

EGP Exterior Gateway Protocol. 70

ETSI European Telecommunications Standards Institute. 41

FTTH Fiber To The Home. 111

IETF Internet Engineering Task Force. 19

IGP Interior Gateway Protocol. 70

IMS IP Multimedia Subsystem. 41

IPTV Internet Protocol television. 41

IS-IS Intermediate System to Intermediate System. 70

NASS Network Attachment SubSystem. 41

NGN Next Generation Network. 39

OSI Open Systems Interconnection. 23

OSPF Open Shortest Path First. 70

- P2P** Peer-to-Peer. 20
- P4P** A Portal for Proactive network Provider Participation for P2P. 24
- PCC** Policies and Charging Control. 45
- PID** Provider-defined network location IDentifiers. 27
- PS** Pair de Stabilité. 114
- QoE** Quality-of-Experience. 17
- QoS** Quality-of-Service. 17
- RACS** Resource and Admission Control Subsystem. 41
- RR** Request Routing. 72
- SQP** Sequential Quadratic Programming. 94
- TE** Traffic Engineering. 38
- TISPAN** Telecommunications and Internet converged Services and Protocols for Advanced Networking. 41

Bibliographie

- [3GP12] 3gpp website, 2012.
- [ABB⁺10] L. Abeni, A. Bakay, M. Biazzini, R. Birke, E. Leonardi, R. Lo Cigno, C. Kiraly, M. Mellia, S. Niccolini, J. Seedorf, T. Szemethy, and G. Tropea. Network friendly p2p-tv : The napa-wine approach. In *Peer-to-Peer Computing (P2P), 2010 IEEE Tenth International Conference on*, pages 1–2, aug. 2010.
- [AK01] E. Altman and H. Kameda. Equilibria for multiclass routing in multi-agent networks. In *Decision and Control, 2001. Proceedings of the 40th IEEE Conference on*, volume 1, pages 604–609 vol.1, 2001.
- [ALT12] Application-layer traffic optimization (alto) website, 2012.
- [APY12] R. Alimi, R. Penno, and R. Y. Yang. Alto protocol. draft 13, IETF, September 2012.
- [Bak05] S. Bakiras. Approximate server selection algorithms in content distribution networks. In *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, volume 3, pages 1490–1494, Mai 2005.
- [Bar08] Michael Bartholomew & “Biggs. Sequential quadratic programming. In *Nonlinear Optimization with Engineering Applications*, volume 19 of *Springer Optimization and Its Applications*, pages 1–14. Springer US, 2008.
- [Bay00] *Théorie des files d’attentes*. Hermes Sciences publications, 2000.
- [BCMP75] Forest Baskett, K. Mani Chandy, Richard R. Muntz, and Fernando G. Palacios. Open, closed, and mixed networks of queues with different classes of customers. *J. ACM*, 22(2) :248–260, April 1975.
- [BJZP09] Yun Bai, Bo Jia, Jixiang Zhang, and Qiangguo Pu. An efficient load balancing technology in cdn. In *Proceedings of the 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery - Volume 07, FSKD ’09*, pages 510–514, Washington, DC, USA, 2009. IEEE Computer Society.
- [BLM⁺11] R. Birke, E. Leonardi, M. Mellia, A. Bakay, T. Szemethy, C. Kiraly, R.L. Cigno, F. Mathieu, L. Muscariello, S. Niccolini, J. Seedorf, and G. Tropea. Architecture of a network-aware p2p-tv application : the

-
- napa-wine approach. *Communications Magazine, IEEE*, 49(6) :154 – 163, june 2011.
- [BM11a] Selim Ellouze Yannick Carlinet Pierre Paris Valery Bastide Emile Stephan Raul Landa Eleni Mykoniati David Griffin Miguel Rio Nico Schwan Klaus Satzke Toufik Ahmed Samir Medjiah Abbas Bradai Ubaid Abbasi Noam Amram Oriol Ribera Prats Bertrand Mathieu, Irène Grosclaude. Initial specification of the envision interface, network monitoring and network optimisation functions. Technical report, 2011.
- [BM11b] Selim Ellouze Yannick Carlinet Pierre Paris Valery Bastide Emile Stephan Raul Landa Eleni Mykoniati David Griffin Miguel Rio Nico Schwan Klaus Satzke Toufik Ahmed Samir Medjiah Abbas Bradai Ubaid Abbasi Noam Amram Oriol Ribera Prats Bertrand Mathieu, Irène Grosclaude. Refined specification of the envision interface, network monitoring and network optimisation functions. Technical report, 2011.
- [BP08] Nicholas Ball and Peter Pietzuch. Distributed content delivery using load-aware network coordinates. In *Proceedings of the 2008 ACM CoNEXT Conference, CoNEXT '08*, pages 77 :1–77 :6, New York, NY, USA, 2008. ACM.
- [BT95] Paul T. Boggs and Jon W. Tolle. Sequential quadratic programming. *Acta Numerica*, 4 :1–51, 1995.
- [BTK94] Paul T. Boggs, Jon W. Tolle, and Anthony J. Kearsley. A truncated sqp algorithm for large scale nonlinear programming problems. In *Advances in Optimization and Numerical Analysis : Proceedings of the Sixth Workshop on Optimization and Numerical Analysis*, pages 69–77. Kluwer Academic Publishers, 1994.
- [CAI]
- [CE10] Niklas Carlsson and Derek L. Eager. Server selection in large-scale video-on-demand systems. *ACM Trans. Multimedia Comput. Commun. Appl.*, 6(1) :1 :1–1 :26, February 2010.
- [Cis05] Cisco. Ospf design guide. Technical report, 2005.
- [Cis13] Cisco. Global mobile data traffic forecast update, 2012-2013. White paper, Cisco, 2013.
- [Coh03] Bram Cohen. Incentives build robustness in bittorrent, 2003.
- [CRM⁺10] Jeff Cohen, Thomas Repantis, Sean McDermott, Scott Smith, and Joel Wein. Keeping track of 70,000+ servers : The akamai query system. Technical report, Akamai, 2010.
- [dEeP13] Rede Nacional de Ensino e Pesquisa. Mapa do backbone, 2013.
- [DOL11] J. Dad, M. Ouali, and Y. Lebbah. A multiclass bcmp queueing modeling and simulation-based road traffic flow analysis. *World Academy of Science, Engineering and Technology*, page 78, 2011.

-
- [Ell12] *Media Networks : Architectures, Applications, and Standards*, chapter Collaboration Between Networks and Applications in the Future Internet. CRC Press, 2012.
- [EMT12] Selim Ellouze, Bertrand Mathieu, and Lemlouma Tayeb. A proposal for improving network-aided cdn server selection. *NOTERE/CFIP 2012*, October 2012.
- [EMT13] Selim Ellouze, Bertrand Mathieu, and Lemlouma Tayeb. A bidirectional network collaboration interface for cdns and clouds services traffic optimization. *IEEE ICC 2013*, June 2013.
- [ETS02] ETSI. Open Service Access (OSA) Application Programming Interface (API) Part 10 : Connectivity Manager SCF. Technical report, ETSI, 2002.
- [FB07] Pierre Francois and Olivier Bonaventure. Avoiding transient loops during the convergence of link-state routing protocols. *IEEE/ACM Trans. Netw.*, 15(6) :1280–1292, December 2007.
- [Fie00] Roy Thomas Fielding. *Architectural Styles and the Design of Network-based Software Architectures*. PhD thesis, UNIVERSITY OF CALIFORNIA, IRVINE, 2000.
- [FPS⁺12] Benjamin Frank, Ingmar Poesse, Georgios Smaragdakis, Steve Uhlig, and Anja Feldmann. Content-aware traffic engineering. *SIGMETRICS Perform. Eval. Rev.*, 40(1) :413–414, June 2012.
- [FT02] B. Fortz and M. Thorup. Optimizing ospf/is-is weights in a changing world. *Selected Areas in Communications, IEEE Journal on*, 20(4) :756–767, may 2002.
- [Gil12] Douglas Gilstrap. Traffic and market report. Technical report, Ericsson, Juin 2012.
- [GK77] M. Gerla and L. Kleinrock. On the topological design of distributed computer networks. *Communications, IEEE Transactions on*, 25 :48–60, 1977.
- [GLP⁺09] C. Griffiths, J. Livingood, L. Popkin, R. Woundy, and Y. Yand. Comcast’s isp experiences in a proactive network provider participation for p2p (p4p) technical trial. Technical report, IETF, 2009.
- [GPSH12] Steffen Gebert, Rastin Pries, Daniel Schlosser, and Klaus Heck. Internet access traffic measurement and analysis. In *Proceedings of the 4th international conference on Traffic Monitoring and Analysis*, TMA’12, pages 29–42, Berlin, Heidelberg, 2012. Springer-Verlag.
- [GSM12] GSMA. Open Network Enablers APIs. Technical report, GSMA, 2012.
- [HAOF05] Y. Higashi, S. Ata, I. Oka, and C. Fujiwara. Topology-aware server selection method for dynamic parallel downloading. In *Consumer Communications and Networking Conference, 2005. CCNC. 2005 Second IEEE*, pages 325 – 330, jan. 2005.

-
- [Her10] Olivier Hersent. *The 3GPP IP Multimedia Subsystem (IMS) Architecture*, pages 269–311. John Wiley and Sons, Ltd, 2010.
- [JBA05] R. Jain, J.L. Bakker, and F. Anjum. *Programming converged networks : call control in Java, XML, and Parlay/OSA*. Wiley, 2005.
- [JG05] Ravi Jain and Xia Gao. *APIs and Application Platforms for Next-Generation Mobile Networks*, pages 157–188. John and Sons, Ltd, 2005.
- [JZSRC09] Wenjie Jiang, Rui Zhang-Shen, Jennifer Rexford, and Mung Chiang. Cooperative content distribution and traffic engineering in an isp network. In *Proceedings of the eleventh international joint conference on Measurement and modeling of computer systems*, SIGMETRICS '09, pages 239–250, New York, NY, USA, 2009. ACM.
- [Kia10] Kiavash Kianfar. *Branch-and-Bound Algorithms*. John Wiley and Sons, Inc., 2010.
- [Kle76] *Queueing Systems, Computer Applications*, volume 2. Wiley, 1976.
- [KMK⁺10] N. Kamiyama, T. Mori, R. Kawahara, S. Harada, and H. Hasegawa. Analyzing influence of network topology on designing isp-operated cdn. In *Telecommunications Network Strategy and Planning Symposium (NETWORKS), 2010 14th International*, pages 1–6, sept. 2010.
- [Kro11] Bill Krogfoss. Analysis : Content peering and the internet economy. Technical report, Alcatel Lucent, 2011.
- [Kuh82] Harold W. Kuhn. Nonlinear programming : a historical view. *SIGMAP Bull.*, (31) :6–18, June 1982.
- [LG12] Kai Lee and GuangYao.Jian. Alto and decade service trial within china telecom. Technical report, IETF, 2012.
- [LIJM09] Craig Labovitz, S. Iekel-Johnson, and D. McPherson. Atlas internet observatory 2009 annual report. Technical report, North American Network Operators' Group, 2009.
- [LIJM⁺10] Craig Labovitz, S. Iekel-Johnson, D. McPherson, J. Oberheide, and F. Jahanian. Internet traffic and content consolidation. Technical report, IETF, 2010.
- [Liu02] Enjie Liu. *A Hybrid Queueing Model for Fast Broadband Networking Simulation*. PhD thesis, Queen Mary, University of London, 2002.
- [LLOH05] Samson Lee, John Leaney, Tim O'Neill, and Mark Hunter. Open service access for qos control in next generation networks – improving the osa/parlay connectivity manager. In Thomas Magedanz, EdmundoR.M. Madeira, and Petre Dini, editors, *Operations and Management in IP-Based Networks*, volume 3751 of *Lecture Notes in Computer Science*, pages 29–38. Springer Berlin Heidelberg, 2005.
- [LMC⁺12] R. Landa, E. Mykoniati, R. G. Clegg, D. Griffin, and M. Rio. Modelling the tradeoffs in overlay-isp cooperation. *International conference of the IFIP Technical Committee on Communication Systems*, 2012.

-
- [LW81] S.S. Lam and J. W. Wong. Queueing network models of packet switching networks. Research report cs-81-06, 1981.
- [MA04] R. T. Marler and J. S. Arora. Survey of multi-objective optimization methods for engineering. *Structural and Multidisciplinary Optimization*, 26(6) :369–395, April 2004.
- [MAB⁺08] Nick McKeown, Tom Anderson, Hari Balakrishnan, Guru Parulkar, Larry Peterson, Jennifer Rexford, Scott Shenker, and Jonathan Turner. Openflow : Enabling innovation in campus networks. Technical report, Openflow, 2008.
- [MDJ11] B. Meskill, A. Davy, and B. Jennings. Server selection and admission control for ip-based video on demand using available bandwidth estimation. In *Local Computer Networks (LCN), 2011 IEEE 36th Conference on*, pages 255 –258, oct. 2011.
- [MES⁺11a] B. Mathieu, S. Ellouze, N. Schwan, D. Griffin, E. Mykoniati, T. Ahmed, and O.R. Prats. Improving end-to-end qoe via close cooperation between applications and isps. *Communications Magazine, IEEE*, 49(3) :136 – 143, march 2011.
- [MES⁺11b] B. Mathieu, S. Ellouze, N. Schwan, D. Griffin, E. Mykoniati, T. Ahmed, and O.R. Prats. Improving end-to-end qoe via close cooperation between applications and isps. *Communications Magazine, IEEE*, 49(3) :136 – 143, march 2011.
- [MK03] A.-J. Moerdijk and L. Klostermann. Opening the networks with parlay/osa : standards and aspects behind the apis. *Network, IEEE*, 17(3) :58 – 64, may-june 2003.
- [MWW08] Jochen Mundinger, Richard Weber, and Gideon Weiss. Optimal scheduling of peer-to-peer file dissemination. *J. of Scheduling*, 11(2) :105–120, April 2008.
- [Nai98] Philippe Nain. Basic elements of queueing theory - application to the modelling of computer systems, 1998.
- [NJWB⁺12] B. Niven-Jenkins, G. Watson, N. Bitar, J. Medved, and S. Previdi. Use cases for alto within cdns. draft 03, IETF, 2012.
- [NSS10] Erik Nygren, Ramesh K. Sitaraman, and Jennifer Sun. The akamai network : a platform for high-performance internet applications. *SIGOPS Oper. Syst. Rev.*, 44(3) :2–19, August 2010.
- [OMA12] Open mobile alliance (oma) website, 2012.
- [Ope12] Open Mobile Alliance. OneAPI Profile of RESTful Network APIs. Technical report, OMA, 2012.
- [Par]
- [PFA⁺10] Ingmar Poese, Benjamin Frank, Bernhard Ager, Georgios Smaragdakis, and Anja Feldmann. Improving content delivery using provider-aided

-
- distance information. In *Proceedings of the 10th ACM SIGCOMM conference on Internet measurement*, IMC '10, pages 22–34, New York, NY, USA, 2010. ACM.
- [PMKN04] Miikka Poikselka, Georg Mayer, Hisham Khartabil, and Aki Niemi. John Wiley and Sons, Ltd, 2004.
- [RFB01] K. Ramakrishnan, S. Floyd, and D. Black. The addition of explicit congestion notification (ecn) to ip. Technical report, IETF, 2001.
- [rGPP09] 3rd Generation Partnership Project. Technical Specification Group Core Network and Terminals; Open Service Access (OSA); Parlay X Web Services; Part 1 : Common (Release 9). Technical report, 3GPP, 2009.
- [rGPP12] 3rd Generation Partnership Project. Technical Specification Group Services and System Aspects; Policy and charging control architecture (Release 11). Technical report, 3GPP, 2012.
- [RH12] Linda Cain Ryan Heath. Digital agenda : Commission opens public consultation on preservation of the open internet (net neutrality). Press release, EUROPEAN COMMISSION, 2012.
- [RHKS02] S. Ratnasamy, M. Handley, R. Karp, and S. Shenker. Topologically-aware overlay construction and server selection. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 3, pages 1190 – 1199 vol.3, 2002.
- [San11] Sandvine. Global internet phenomena spotlight - netflix rising. Technical report, Sandvine Incorporated, 2011.
- [SB09] J. Seedorf and E. Burger. Application-layer traffic optimization (alto) problem statement. RFC 5693, IETF, Octobre 2009.
- [Sch86] K. Schittkowski. Nlpql : A fortran subroutine solving constrained nonlinear programming problems. *Annals of Operations Research*, 5 :485–500, 1986.
- [SE12] Emile Stephan and Selim Ellouze. Alto session for cdn interconnection draft-stephan-cdni-alto-session-ext-01, 2012.
- [TcSfANT09a] Telecommunications, Internet converged Services, and Protocols for Advanced Networking (TISPAN). NGN Functional Architecture. Technical report, TISPAN, 2009.
- [TcSfANT09b] Telecommunications, Internet converged Services, and Protocols for Advanced Networking (TISPAN). NGN Functional Architecture ; Resource and Admission Control Subsystem (RACS). Technical report, TISPAN, 2009.
- [TcSfANT10] Telecommunications, Internet converged Services, and Protocols for Advanced Networking (TISPAN). NGN Functional Architecture ; Network Attachment Sub-System (NASS). Technical report, TISPAN, 2010.

- [TIS12] Tispan - defining the next generation network, 2012.
- [UCoL09] Alcatel-Lucent LABRI LiveU Telefonica I+D University College of London, Orange Labs. Envision : Co-optimisation of overlay applications and underlying networks, 2009.
- [Unm11] Musa Unmehopa. Supporting innovation in an open ict service environment. In *ICIN 2011, From Bits to Data, From Pipes to Cloud*, 2011.
- [UVB06] Musa Unmehopa, Kumar Vemuri, and Andy Bennett. *Part Introduction*, pages 87–87. John Wiley and Sons, Ltd, 2006.
- [WJFR10] Patrick Wendell, Joe Wenjie Jiang, Michael J. Freedman, and Jennifer Rexford. Donar : decentralized server selection for cloud services. *SIGCOMM Comput. Commun. Rev.*, 41(4) :-, August 2010.
- [WLCM11] Jianxiong Wan, Chuang Lin, Xin Chen, and Kun Meng. On the optimal request routing strategy in cdn live streaming application. In *Communications (ICC), 2011 IEEE International Conference on*, pages 1 –5, june 2011.
- [XKSY] Haiyong Xie, Arvind Krishnamurthy, Avi Silberschatz, and Richard Y. Yang. P4p : Explicit communications for cooperative control between p2p and network providers.
- [XYK⁺08] Haiyong Xie, Y. Richard Yang, Arvind Krishnamurthy, Yanbin Grace Liu, and Abraham Silberschatz. P4p : provider portal for applications. *SIGCOMM Comput. Commun. Rev.*, 38(4) :351–362, August 2008.
- [Zad63] L. Zadeh. Optimality and non-scalar-valued performance criteria. *Automatic Control, IEEE Transactions on*, 8(1) :59 – 60, jan 1963.