



Architectures des réseaux pour le contrôle de la QoS

Olivier Dugeon

► To cite this version:

Olivier Dugeon. Architectures des réseaux pour le contrôle de la QoS. Réseaux et télécommunications [cs.NI]. Institut National Polytechnique (Toulouse) (Toulouse INP), FRA., 2008. <tel-04321775>

HAL Id: tel-04321775

<https://hal.science/tel-04321775v1>

Submitted on 4 Dec 2023

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

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



HAL Authorization

Mémoire d'Habilitation à Diriger des Recherches

Architectures des réseaux pour le contrôle de la QoS

Olivier Dugeon

Ingénieur Chercheur à Orange Labs

présentée et soutenue publiquement le 18 décembre 2008

Composition du jury

Président:	Richard Castanet, Professeur à l'Université de Bordeaux
Rapporteurs:	Anna Cavalli, Professeur à l'INT d'Evry Béatrice Paillassa, Professeur à l'INP de Toulouse Edmundo Monteiro, Professeur à l'Université de Coimbra
Correspondant:	Michel Diaz, Directeur de Recherche au LAAS-CNRS
Examineurs:	Jean-Marie Farines, Professeur à l'UFSC

*A Sylvie,
Marine,
et Julie*

Je remercie chaleureusement Michel Diaz pour m'avoir suivi et conseillé judicieusement dans l'élaboration et la rédaction de ce mémoire d'Habilitation à Diriger des Recherches.

Je remercie Anna Cavalli, Béatrice Paillassa et Edmundo Monteiro d'avoir accepté de rapporter cette Habilitation à Diriger des Recherches ainsi que Jean-Marie Farines, Richard Castanet et Michel Diaz d'avoir accepté d'être membres du Jury.

Pour m'avoir donné la possibilité de mener mes travaux de recherches et favoriser leur reconnaissance en préparant cette Habilitation à Diriger des Recherches, je remercie toutes les personnes et collègues d'Orange Labs avec lesquels j'ai collaboré.

Je tiens également à exprimer ma gratitude envers Sylvie, Marine et Julie à qui je dédie ce mémoire pour leur aide morale et matérielle.

Résumé

La qualité de service dans les réseaux téléphoniques a toujours tenu une place importante, voire prépondérante, dans la conception des architectures de télécommunication. Si elle est implicitement supportée dans le réseau téléphonique commuté (RTC) classique, il n'en est pas de même dans les réseaux en mode paquet que sont l'ATM et l'IP. Ce mémoire d'habilitation à diriger des recherches synthétise le travail de recherche effectué sur un ensemble d'architectures pour contrôler la QoS dans les réseaux en mode paquets. De l'espace dans le plan de transfert, jusqu'à la négociation des contrats avec l'utilisateur en passant par le plan de contrôle, les contributions forment à la fois une continuité et une complémentarité permettant de dégager des règles d'urbanisme à suivre pour la gestion de la QoS. A partir de ces règles, une solution globale de gestion de la QoS depuis le réseau local jusqu'à l'inter-domaine est proposée. Elle suit également une découpe temporelle séparant ce qui procède de l'approvisionnement du réseau et de l'invocation permettant à la solution d'être extensible et applicable à grande échelle et dont les fondements principaux sont :

- Une gestion des files d'attente et du multiplexage par espacement comme discipline de service universel garantissant un délai, une gigue et un taux de perte faible tout en respectant la caractéristique temporelle intrinsèque de chaque application,
- Une gestion par classes de services approvisionnées par MPLS-DiffServ-TE à l'intra et à l'inter domaine durant le processus de commande / livraison,
- Une admission d'appel pour protéger les ressources affectées aux classes de services durant le processus d'invocation,
- Et une intégration de la gestion de la QoS multiservices dans une architecture multi technologies.

Mots-clés : Qualité de Service, Classe de Service, architecture, urbanisme, médiation, espacement, admission d'appel, ingénierie de trafic, approvisionnement, ATM, IP.

Abstract

The network Quality of Service has already taken an important place in the design of telecommunication architectures. If QoS is implicit in classic telephony network (PSTN), it is not the same in packet networks like ATM and IP. This memory of "Habilitation à Diriger des Recherches" synthesis the research studies done to various architectures in order to provide quality of service in packet networks. From spacing in transfer plane to QoS negotiation with user through the control plan, the contributions are complementary and form a continuity suitable to define urban networking rules to manage the quality of service. From these urban networking rules, a global solution to manage the QoS from the local network to the interdomain is proposed. It is splits between the provisioning phase and the invocation phase allowing the solution to be scalable. The main features are:

- A spacer queuing and multiplexing scheme as a universal queuing service which guarantee a low delay, jitter and loss rate with respect to the intrinsic temporal characteristic of each application,
- A per class of service management provisioned by MPLS-DiffServ-TE at both intra and inter domain during the provisioning process,
- An admission control to protect the resources allocated to each Class of Services during the invocation process,
- And an integration of the QoS control in a multi-service and multi-technology architecture.

Keywords : Quality of Service, Class of Service, architecture, urban networking, mediator, spacer, call admission control, traffic engineering, provisioning, ATM, IP.

Table des matières

Introduction et CV	1
1. Contexte des travaux de recherches	2
2. Curriculum Vitæ	4
2.1. État Civil	4
2.2. Diplômes	4
2.3. Activités de recherches	4
2.4. Travaux actuels	6
3. Activités pédagogiques	7
3.1. Encadrements	7
3.2. Enseignements	8
3.3. Lecteur – Correcteur – Revue	8
3.4. Diffusion des connaissances	8
4. Bibliographie	9
4.1. Journaux avec comité de lecture	9
4.2. Participation à ouvrage	9
4.3. Conférences avec actes et Comité de lecture	9
4.4. Brevets	10
4.5. Contributions aux standards	10
4.6. Rapports de projets européens	11
4.7. Rapports internes	11
4.8. Présentations dans des conférences internationales	12
Architectures de gestion de la QoS	13
1. Définitions	14
1.1. La qualité de service	14
1.2. Paramètres de QoS	14
1.3. Les classes de services	15
1.4. Les ressources	16
1.5. Les contrats	16
1.6. Flot, micro flot, agrégat et session	17
2. Fonctions élémentaires de gestion de la QoS	18
2.1. Décomposition des équipements	18
2.2. Les différentes méthodes de gestion de la QoS	19
2.3. Les points de congestions	20
2.4. Principes de base à respecter	21
3. Outil de représentation des architectures	22
3.1. Le modèle OURAf	22
3.2. Principe	22

3.3. Les niveaux	22
3.4. Les Processus	24
3.5. Les partitions	26
3.6. Représentation fonctionnelle	26
4. La commande des ressources	28
4.1. Application de OURAf	28
4.2. Approvisionnement	29
4.3. Invocation	30
4.4. Surveillance et Mesure	31
5. L'admission d'appel	33
5.1. Définition	33
5.2. CAC de niveau service	34
5.3. CAC de niveau réseau	35
5.4. CAC de niveau transfert	35
6. Les architectures standardisées	36
6.1. Le modèle ATM	36
6.2. Les modèles de l'IETF	38
6.3. Les modèles du NGN	45
7. Analyses des solutions	48
7.1. Synthèse de l'état de l'art	48
7.2. Limitations des solutions actuelles	49
7.3. QoS mythe ou réalité ?	50
7.4. Conclusions	53
Contributions aux travaux d'Architectures pour la QoS	55
1. Contrôle et Multiplexage par espacement	56
1.1. Espacement et échéancier virtuel	56
1.2. Le multiplexage par espacement	58
1.3. Régulation de TCP par espacement	59
1.4. Conclusion	62
2. La signalisation allégée	63
2.1. Présentation du projet ASIA	63
2.2. Couplage ABT et MPLS	63
2.3. Partage des ressources avec l'ABT/IT	63
2.4. Synchronisation de TCP avec ABT/IT	64
2.5. Conclusion et bilan	65
3. La médiation d'accès aux ressources	67
3.1. Présentation du projet Cadenus	67
3.2. Architecture basée sur la médiation	67
3.3. Du SLA au SLS	68

3.4. Architecture du Resource Mediator	70
3.5. Le Contrôleur de Réseaux	71
3.6. Modèle « Hub » versus « Cascade »	71
3.7. Conclusion et Bilan	72
4. Service de QoS à la demande	73
4.1. Description	73
4.2. Limitations du Flow Spec	75
4.3. Sécurité, authentification et comptage	75
4.4. Applications	76
5. Admission d'appel à l'inter-domaine	77
5.1. Architecture EuQoS	77
5.2. Approvisionnement et utilisation des chemins à QoS	78
5.3. Construction des EQ-Path selon le « hard model »	81
5.4. Bilan et Conclusion	84
6. Acquisition de topologie	85
6.1. STAMP	85
6.2. Agrégation de topologie	86
6.3. Conclusion	87
7. Admission d'appel dans le réseau domestique	89
7.1. La passerelle résidentielle	89
7.2. Admission d'appel répartie	92
Bilan et Perspectives	95
1. Bilan des travaux	96
1.1. Synthèse	96
1.2. Règles d'urbanisme	97
1.3. Architecture de référence	99
2. Perspectives	104
2.1. Le routage hiérarchique	104
2.2. Évolution des algorithmes de CAC pour les réseaux domestiques	106
2.3. La préemption et les interactions de service	107
2.4. Le partage multimédia	108
Bibliographie	109

I

Introduction et CV

« Inventer, c'est penser à côté »

Albert Einstein

Ce mémoire d'habilitation a pour thème les architectures des réseaux pour la gestion et le contrôle de la qualité de service. Il ne vise pas à détailler des technologies particulières mais analyse le problème de la garantie de la qualité de service dans sa globalité y compris de bout en bout, et présente les apports de l'auteur dans le domaine. Aussi, le lecteur est invité à se rapporter aux références bibliographiques pour les détails de l'état de l'art et les contributions techniques. Le mémoire est organisé comme suit.

Après cette partie introductive, la deuxième partie inventorie les technologies de contrôle de la qualité de service dans les réseaux en mode paquet à l'aide d'un outil de représentation des architectures fonctionnelles. Suit alors une partie exposant nos contributions dans le domaine des technologies et des architectures pour le contrôle de la qualité de service à l'aide de fonctions d'admission d'appels, tout d'abord dans le plan de transfert puis dans le plan de contrôle. Enfin, la dernière partie conclut et dresse un bilan de l'adoption par les opérateurs des mécanismes de gestion de la qualité de service en tentant de répondre à la question : « QoS mythe ou réalité ? » puis en exposant les recommandations de l'auteur en la matière, ainsi que les études restant à mener pour finaliser la solution préconisée.

Cette première partie définit le cheminement et la continuité des études menées dans le domaine du contrôle d'admission dans les réseaux en mode paquet. Le deuxième chapitre présente mon curriculum vitae détaillé et fournit des renseignements d'ordres administratifs, ainsi qu'un résumé de mon parcours professionnel et de mes travaux de recherches. Le chapitre suivant décrit mes activités pédagogiques et d'encadrements avant de terminer par ma bibliographie.

1. Contexte des travaux de recherches

La qualité de service dans les réseaux téléphoniques a toujours tenu une place importante voire prépondérante dans la conception des architectures de réseaux. Si elle est implicitement supportée dans le réseau téléphonique commuté (RTC) classique, il n'en est pas de même dans les réseaux en mode paquet que sont l'ATM et l'IP. Les réseaux informatiques, quant à eux, n'ont jamais placé la QoS au centre de leur préoccupation. A la fin des années 80, le besoin s'est fait sentir de disposer d'un réseau universel à même de transporter tout type de services. Des différentes technologies en compétition, l'ATM (Asynchronous Transfert Mode) sera standardisée comme la technologie réseau multiservice. Très vite, les travaux de recherche ont démarré afin de garantir la qualité de service. En effet, d'un mode de transport circuit garantissant de la QoS mais uniquement pour un service (la téléphonie) et un débit (64 kbit/s), l'ATM a libéré deux dimensions dans la flexibilité du réseau : le multi débit et le multiservice. Les problèmes à résoudre étaient vastes et complexes. De nombreuses solutions ont été élaborées allant des nouveaux modèles mathématiques de modélisation du multiplexage aux nouvelles disciplines de file d'attente. En même temps, passant des réseaux locaux et de leurs interconnexions, les réseaux informatiques et en particulier la technologie IP se sont peu à peu imposés comme réseau de transport universel. Cependant, en comparaison avec l'ATM, ces deux nouveaux degrés de libertés qui ont été libérés : les paquets sont à longueur variable et l'acheminement s'effectue en mode non connecté. Ajoutées à un mode de transfert non fiable, ces deux dimensions supplémentaires dans la flexibilité des réseaux sont devenues un véritable casse-tête pour la gestion de la QoS. Outre ces considérations techniques, c'est peut-être avant tout un choc culturel entre le monde informatique et le monde des télécommunications qui a le plus porté préjudice aux études sur la QoS. Entre l'incompréhension des premiers et le jusqu'au-boutiste des seconds, il aura fallu attendre une bonne décennie avant qu'un consensus s'établisse sur les impératifs et les besoins de la qualité de service pour nombre d'applications qui n'avaient pas été prévues à la création du protocole IP. Que ce soit pour la VoIP ou l'IP TV, il est maintenant indispensable pour les opérateurs de pouvoir disposer d'un contrôle de la qualité de service. C'est en effet elle qui conditionne en grande partie le succès d'une offre sur le long terme.

Mais la Qualité de Service (QoS en anglais) couvre un large spectre allant de la sûreté de fonctionnement des équipements à la gestion optimale des ressources du réseau. C'est à cette dernière que se réfère ce mémoire d'habilitation. En effet, la garantie de QoS au niveau du réseau, vis-à-vis de la demande d'un service, doit mettre en œuvre de nombreux mécanismes adaptés et cohérents entre eux. Ainsi, elle est assurée d'une part par la protection du réseau contre les surcharges (fonction d'Admission d'Appel), par la maîtrise des chemins (fonction de Réservation de Ressources) et par la configuration adéquate des équipements du réseau traversés par les données du service (fonction de Commande de Ressources). Ce travail d'habilitation à diriger des recherches correspond à la synthèse de mes travaux menés dans ce domaine au cours de ces 10 dernières années, en montrant l'évolution de la réflexion, du cheminement des idées, l'analyse des différents résultats obtenus ainsi que la présentation d'une solution globale. Ces travaux ont été conduits dans un contexte de recherche opérationnel avec la contrainte de trouver des solutions adaptées aux réseaux d'opérateurs de télécommunications.

Au fil des différents postes et des études menées, une thématique cohérente apparaît : Comment fournir de la qualité de service (QoS) à l'aide de mécanismes d'admission d'appel et de commande de ressources dans les réseaux de transfert en mode paquet ? Démarrées avec l'ATM puis progressivement migrant vers l'IP, mes études ont également évolué en

fonction des besoins mais surtout en s'adaptant à l'évolution des architectures réseaux. On peut cependant les classer suivant une découpe en niveaux.

Ces travaux ont commencé par le niveau paquet. Ayant débuté par l'étude des fonctions de multiplexage statistique puis d'espacement (appliquées aux fonctions de police et de multiplexage), la première évolution s'est effectuée en passant du niveau cellule ATM au niveau trame via la réservation rapide de ressources pour des blocs de cellules. Toujours en restant en mode paquet, l'évolution vers IP a commencé en regardant comment les mécanismes développés pour l'ATM pouvaient être transposés à la technologie IP. L'admission d'appel décrit ici l'ensemble des fonctions de polices appliquées au niveau paquet, mais également au niveau trame. La réservation de ressources désigne soit l'occupation du lien de sortie d'un multiplexeur, commutateur ou routeur, soit la réservation de ressources de bout en bout à l'aide de cellules spécifiques de réservation de ressources. Cette période a été synthétisée à travers ma thèse de 3^{ème} cycle soutenue en 1997 et ayant pour titre « Machine d'Admission pour le Réseau d'Infrastructure ATM ». Un brevet a été déposé durant cette même période. 4 publications ont également été présentées et publiées.

Ensuite, le besoin de coupler les mécanismes de niveau paquet avec les protocoles des couches supérieures s'est rapidement imposé. Cette période correspond aux études de signalisation allégée et de leur application dans les architectures NGN qui préconisent la séparation des couches de services, de transports et de contrôles. La charnière avec la période précédente s'est effectuée avec les études sur l'espacement de TCP et sur les mécanismes en couche ATM permettant de réguler TCP. Au fur et à mesure de l'étude du couplage de TCP avec les techniques de réservation ATM, il est apparu que ce couplage devait s'effectuer plus haut dans les couches, i.e. au niveau du service. En même temps les prémisses du besoin d'informations situées au niveau transport pour le plan de contrôle se sont fait ressentir. Malgré la séparation en couche préconisée par le modèle NGN, le plan de contrôle a besoin de rester synchronisé non seulement avec le plan de transfert, mais avec le plan de service. Cette période est synthétisée dans la préparation et la participation au projet RNRT ASIA. Un brevet a été déposé à cette occasion et 8 publications ont parachevé le travail.

Finalement, en considérant l'architecture globale et les aspects multiservices, les études sur l'admission d'appel et la commande de ressources ont porté sur la création d'un service de QoS de bout en bout pour les offres multiservice (Voix, Vidéo, Internet) et l'étude des plates-formes de commande de ressources. Plus orientés vers l'architecture des réseaux, ces travaux ont permis de mieux cerner la problématique de la gestion de la QoS dans son ensemble et non plus morceau par morceau. En partant de l'application aux réseaux d'accès du type xDSL, mes travaux se sont progressivement étendus du réseau cœur au réseau domestique en prenant en compte l'inter-domaine pour aboutir à une solution globale de la gestion de la QoS de bout en bout. Cette période correspond à la mise en place et la conduite de 2 projets internes (CER puis QoSTicS) et de deux projets européens (CADENUS puis EuQoS) ainsi qu'à l'encadrement d'une thèse. Ces travaux ont donné lieu au dépôt de deux brevets sur les algorithmes d'admission d'appel et la publication de 12 articles.

2. Curriculum Vitæ

2.1. État Civil

Olivier Dugeon, né le 10 Avril 1966 à Paimboeuf (Loire-Atlantique) nationalité Française

Ingénieur Chercheur à Orange Labs (France Telecom R&D)

2, Avenue Pierre Marzin 22307 Lannion Cedex

Tél/Fax: +33 296 052 880 - +33 682 903 785 / +33 296 051 470

olivier.dugeon@orange-ftgroup.com

2.2. Diplômes

- **1984:** Baccalauréat série E
- **1986:** Diplôme Universitaire de Technologie (D.U.T.) Génie Électrique de l'Institut Universitaire de Technologie (I.U.T.) de Lannion
- **1989:** Diplôme d'ingénieur en Électronique et Informatique Industrielle de l'École Nationale Supérieure des Sciences Appliquées et de Technologies (ENSSAT) de Lannion
- **1997:** Doctorat Informatique de l'Université de Rennes I
- **1998:** Qualification par le CNU section 27 (Informatique) aux fonctions de Maître de Conférence

2.3. Activités de recherches

2.3.1. 1990 - 1992

Pendant 3 ans, j'ai été chargé de l'étude et de la réalisation d'un simulateur matériel afin d'accélérer une simulation logicielle. Il a permis l'étude de l'enveloppe de multiplexage d'un ensemble de trafics ATM à débit variable par palier de durée constante. Son rôle était de déterminer le débit et la mémoire nécessaires au support de nouvelles connexions en l'absence d'une boucle de contrôle durant la phase de transfert. Comparée à une version logicielle sur station de travail performante, la machine de simulation fournissait des résultats 10 à 100 fois plus vite et permettait l'analyse des saturations de files d'attente - jusqu'à une probabilité de l'ordre de 10^{-8} . Cette étude fut présentée au RACE 1022, projet Européen auquel j'ai participé durant les années 1991-1992.

La machine de simulation a été ensuite transférée à l'INT d'Evry où une équipe d'enseignants chercheurs a continué son exploitation. Des étudiants de 3^{ème} année, au cours de leurs projets de fin d'année, ont enrichi les outils de constructions des jeux de données à simuler. Cette étude a donné lieu à plusieurs encadrements de stagiaires.

2.3.2. 1993 - 1995

J'ai étudié le principe de multiplexage par échancier de connexions ATM provenant de différentes entrées physiques. Cette étude a donné lieu, fin 1993, au dépôt d'un brevet sur cette nouvelle discipline de file d'attente. J'ai implémenté ce nouveau principe de multiplexage décrit dans le brevet dans des FPGA à partir d'une synthèse VHDL. J'ai présenté à plusieurs reprises une démonstration montrant l'intérêt de la fonction de contrôle et d'espacement lors des événements suivants :

- ATM développement 94 à Rennes,
- INTEROP'94 à Paris,
- 50^{ème} Anniversaire du CNET à Paris en 1995,
- Ainsi que pour illustrer le cours ATM dispensé au CNET Lannion.

2.3.1. 1996 – 1997

Durant ces deux années, outre la rédaction de ma thèse, j'ai étudié le concept de Network Processor pour l'ATM. Une fois identifiées les fonctions propres aux traitements des cellules ATM, j'ai élaboré différentes architectures. Une pré version de cette plate-forme m'a permis de réaliser une démonstration du protocole ABT/IT afin de transporter de façon transparente les protocoles d'Internet (TCP, UDP) sur un réseau ATM longue distance via un service UBR+ à bande passante flexible. J'ai présenté cette démonstration lors des congrès ATM 97 à Tel-Aviv en Israël, ATM Dev'97 et 98 à Rennes. Cette démonstration a également servi à illustrer le cours ATM du CNET.

2.3.2. 1997 – 2000

J'ai participé au montage du projet de signalisation allégée ASIA ("Accelerated Signalling for the Internet over ATM") qui a été labellisé par le RNRT fin 98. J'ai étudié les principes de la signalisation allégée et leur mise en œuvre. J'ai participé à la rédaction du brevet sur le couplage de TCP avec la signalisation allégée ainsi qu'à 4 publications. J'ai réalisé et démontré les principes de signalisation allégée au Directeur de la BRX et à son état major en juin 1999. La version définitive de la maquette ASIA a été présentée également au représentant du ministère de l'industrie. J'ai suivi une Consultation Thématique Informelle (CTI) avec l'Université de RENNES-I et participé au GDR-ARP sur les Network Processor.

2.3.3. 2001 - 2004

Durant cette période, j'ai encadré et contribué au projet de recherche interne de France Telecom R&D sur la commande de ressources et de services, nommé CER (Commandes Et Ressources). Ce projet, d'une durée de 2 ans et d'un volume de 250 hommes mois, réunissait une dizaine de collaborateurs (Ingénieurs chercheurs, techniciens et stagiaires).

Conjointement au projet CER, j'ai assuré l'encadrement technique et la coordination du projet IST FP5 CADENUS conjointement avec Martin Potts (Société Martel). D'une durée de 3 ans, ce projet regroupait 3 opérateurs (France Telecom, Eircom et Telecom Italia), 3 universitaires (EPFL, Université de Naples, Teltec DCU), 2 industriels (Sodalía et GMD Fokus) et 2 consultants (Martel et Telscom), soit une quinzaine de partenaires, pour un budget de 6,5 M€. Les principaux résultats du projet ont été (conjointement au projet CER et CADENUS) :

- Une architecture de référence pour le plan de contrôle des réseaux permettant de mutualiser la commande des différents services des opérateurs (VoD, VoIP, Internet),
- Des règles d'urbanisme à suivre et des préconisations vers les opérateurs afin de déployer une solution de commande mutualisée conforme à cette architecture,
- Une spécification des plates-formes de commande de ressources,
- Un démonstrateur des principes de commandes de ressources multiservices (TriplePlay) présenté devant les membres du comité exécutif de France Telecom et le TOP 100 des managers de France Telecom,
- De nombreuses publications (30) dont l'article lors de la conférence ISS/WTC2002 qui a donné lieu à une démonstration publique du projet CADENUS.

2.3.4. 2004-2007

Durant cette période, j'ai encadré et contribué à un deuxième projet de recherche, nommé QoS_{TicS}, au sein de France Telecom R&D. D'une durée de 2 ans pour un volume de 350 hommes mois, le projet regroupait une équipe de 20 personnes dont une dizaine à temps plein. L'objectif était d'étudier et définir un service de QoS à la demande mutualisé pour les services TriplePlay (VoD, VoiP, Internet). Les principaux résultats ont été :

- Spécification du service de QoS mutualisé à la demande,
- Définition de l'architecture du plan de commande support du service,
- Démonstration des principes sur plate-forme captive,
- Spécification des interfaces,
- Contributions vers les instances de normalisations ETSI et ITU-T qui ont donné lieu à la publication des spécifications du RACS (Resource Allocation Control Subsystem) à l'ETSI et à la recommandation Y.2111 - RACF (Resource Allocation Control Function) à l'ITU.

Conjointement à ce projet, j'ai assuré la co-ordination des contributions de FTR&D au projet européen IST/EuQoS. Les résultats du projet QoS_{TicS} ont servi de base aux contributions de FTR&D au projet IST/EuQoS et ont permis de définir l'architecture du système EuQoS, et plus particulièrement, la définition de l'architecture et la réalisation du système d'approvisionnement de la QoS de bout en bout basé sur MPLS Traffic Engineering.

Durant cette même période, j'ai étudié, spécifié, réalisé puis normalisé les fonctions d'admission d'appel pour les passerelles domestiques (LiveBox) pour le forum Home Gateway Initiative.

2.4. Travaux actuels

Je mène des travaux de recherches dans le domaine de la commande de ressources et l'admission d'appel et plus particulièrement :

Dans le réseau domestique sur:

- L'interaction de la fonction d'admission d'appel multiservices dans la passerelle domestique (Home Gateway) et celle du réseau d'accès pour contrôler les ressources du réseau domestique et du réseau d'accès,
- La mise en place de connexions à QoS garantie comme support d'un service de partage de contenus multimédias entre réseaux domestiques,
- Le contrôle d'admission répartie entre plusieurs passerelles domestiques qui a fait l'objet d'un dépôt de brevet,
- L'interopérabilité des fonctions d'admission d'appel entre les différents standards (UPnP-QoS, DLNA, HGI ...) des réseaux domestiques.

Et dans les réseaux de cœur sur:

- L'approvisionnement de ressources à l'inter-domaine basé sur MPLS-DiffServ Traffic Engineering,
- L'acquisition de topologie pour alimenter les bases de données des plates-formes de contrôle d'admission d'appel (standard RACS de l'ETSI-TISPAN et RACF de l'ITU-T),
- Le routage hiérarchique inter-domaine basé sur l'agrégation de topologie.

3. Activités pédagogiques

3.1. Encadrements

3.1.1. Stages

Je propose des sujets de stage pour des étudiants en BAC+5 (école d'ingénieur) et j'encadre des stagiaires BAC+5 :

- 1 DEA et 4 élèves ingénieurs à hauteur de 100% en 1992, 1993, 1999, 2000 et 2001,
- 3 élèves ingénieurs à hauteur de 50% en 1998, 2002 et 2004.

Et plus récemment, de Février à Août 2007, j'ai encadré à 100% le stage de DEA de Mr Kassan Rabih sur l'intégration d'un serveur de placement de tunnel MPLS (PCE) dans le système EuQoS. De Février à Août 2008, j'ai encadré à 100% le stage de Master de Mr Samuel Rossec sur le développement du service UPnP QoS version 3 sur une passerelle domestique.

3.1.2. Thèse

Dans le cadre du projet RNRT ASIA et du contrat (CTI) entre FTR&D et l'Université de Rennes I, j'ai participé à l'encadrement à hauteur de 25% de la thèse de Ferry Djiéya Nganchui (Voir lettre d'accompagnement de François Charrot). La thèse, portant sur la « Contributions à la conception de systèmes de contrôle de trafic et de gestion de ressources en ATM », a été soutenue le 4 décembre 2002 à l'université de Rennes I.

Au sein du projet CADENUS, j'ai participé à hauteur de 25% à l'encadrement de la thèse d'Ada Diaconescu portant sur la médiation de service. La thèse a été soutenue en 2006. Cet encadrement a donné lieu à l'article sur l'architecture CADENUS présenté lors de la conférence ISS/WTC2002.

Depuis novembre 2005, j'encadre le travail d'un doctorant, Mr Htira Walid, sur la synchronisation du plan de transfert et du plan de contrôle à hauteur de 75%. Michel Diaz du CNRS-LAAS étant le co-encadrant à hauteur de 25% de cette thèse. Les résultats ont donné lieu à cinq publications :

- Sur un nouveau modèle d'approximation de la modélisation de topologie de réseau pour l'admission d'appel à la conférence IEEE Symposium on Computer and Communication (ISCC'07) pour lequel M. Walid Htira a obtenu le « Best Student Paper Award »,
- Sur l'agrégation de topologie de réseau en étoile en fonction du délai et de la bande passante au 12th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD '07),
- Sur l'agrégation de topologie de réseau en modèle « fulls-mesh » à l'Australian Telecommunication Networks and Applications Conference (ATNAC'07).
- Sur un nouveau protocole de découverte automatique de la topologie des réseaux à la 22^{ième} conférence internationale Advanced Information Networking and Applications (AINA'08),
- Enfin sur un nouveau schéma d'admission d'appel basé sur l'agrégation de topologie à la conférence IFIP Networking 2008.

Ainsi qu'à un dépôt de brevet sur une technique de calcul du centre de la topologie agrégée en vue d'optimiser l'admission d'appel à l'INPI sous le numéro 07 07043.

3.2. Enseignements

J'ai assuré les cours suivants :

- **1989 – 1990:** Cours de réseau aux 2^{ème} année (TD, TP) et année spéciale (CM, TD et TP) du département Informatique de l'IUT de Lannion ainsi que les cours de système pour l'année spéciale et 1^{ere} année LSI de l'ENSSAT (CM, TD, TP).
- **1990 – 1993:** Cours réseau aux 2^{ème} année du département Informatique à l'IUT de Lannion
- **1992 – 1993:** Cours du C.N.A.M «circuit numérique» du cycle B.
- **1994 – 2000:** TP de simulation numérique (modélisation des files d'attentes ATM en QNAP), aux étudiants de Master et de 3^{ème} année de l'Institut National des Telecom (INT) d'Evry ainsi qu'au titre de la formation continue.
- **2000 – 2006:** Cours Réseau IP aux étudiants de 2^{ème} année informatique de l'ENSSAT dans le cadre de la formation initiale. Mise en place et TP réseau (configuration de base des routeurs, routage statique et protocole de routage dynamique) lors des 2 dernières années.
- **2007:** Cours Réseau dans le système Linux aux étudiants de première année du Master réseau de l'Université de Brest Occidentale (UBO).

3.3. Lecteur – Correcteur – Revue

Membre du comité de programme (TPC) de la conférence « High Performance Computing Conference (HPCC) »

Relecture de papier pour diverses conférences (Networking 2006, Toward the QoS Internet workshop, NOMS 2006/2007/2008, ICAS'05 & ICNS'05, IARA) et pour l'IEEE communication magazine (5 numéros)

3.4. Diffusion des connaissances

Je diffuse mes connaissances en interne en tant qu'expert auprès des collègues de mon équipe aussi bien qu'auprès des collaborateurs qui contribuent aux projets dont j'ai la responsabilité. Je suis également amené en tant qu'expert à apporter mes connaissances et compétences dans mon domaine dans divers projets au sein de France Telecom et en externe.

J'ai organisé la conférence "First European Workshop on End-to-end QoS in the Internet - EEQoS'2005" à Paris en Juin 2005.

4. Bibliographie

4.1. Journaux avec comité de lecture

1. O. Dugeon, E. Mingozi, G. Stea, L. Bisti, « *Provisioning QoS in Inter-domain traffic Engineering* » - Annals Of Telecommunications, volume 63 – n° 11-12 pages x-y - 2008
2. O. Dugeon. « *Du contrat aux spécifications jusqu'à la commande de ressources: l'architecture CADENUS* » – Revue de l'Electricité et de l'Electronique (REE) n° 10 – Novembre 2002
3. F. Guillemin, O. Dugeon, J. Boyer and C. Mangin. "Lightweight signaling in ATM networks for high quality transfer of Internet traffic" - Computer Networks – volume 34 pages 263-278 - 2000
4. J. Boyer, P. Boyer, O. Dugeon, F. Guillemin and C. Mangin. "Accelerated signalling for the internet over ATM" - European Transactions on Telecommunications and Related Technologies (ETT) – volume 10(2) pages 153-164 – 1999

4.2. Participation à ouvrage

1. T. Braun, M. Diaz, J.E. Gabeiras, T. Staub. « End-to-End Quality of Service Over Heterogeneous Network » - chapitre 6, pages 137–184 – Springer, 2008
2. J. Pellaumail, P. Boyer et P. Leguesdron. « *Autoroutes ATM* » - chapitre 11, pages 213-224 - Hermes, 1997

4.3. Conférences avec actes et Comité de lecture

1. W. Htira, O. Dugeon, and M. Diaz, "A Novel Bandwidth Broker Architecture Based on Topology Aggregation in Delay/Bandwidth Sensitive Networks", IFIP Networking 2008, Singapour, May 2008
2. M. A. Callejo Rodríguez *et al.* "EuQoS: End-to-End QoS over Heterogeneous Network", ITU-T Kaleidoscope Conference – Innovations in NGN, Geneva, 12-13 May 2008.
3. W. Htira, O. Dugeon, and M. Diaz, "STAMP: Towards A Scalable Topology Announcement and Management Protocol", 22nd IEEE International Conference on Advanced Information Networking and Applications (AINA2008), Ginowan, Okinawa, Japan, March 25 - 28, 2008
4. W. Htira, O. Dugeon and M. Diaz. "A Mesh Aggregation Scheme For Call Admission Control", Australasian Telecommunication Networks and Applications Conference (ATNAC'07), New-Zealand, December 2007
5. W. Htira, O. Dugeon and M. Diaz. "An Aggregated Delay|Bandwidth Star Scheme For Admission Control" – 12th International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD '07) – In Proc of 18th Annual IEEE International Symposium on Personal, Indoor and Radio Communications – (PIMRC'07), Athens, September 2007
6. W. Htira, O. Dugeon and M. Diaz. "A New Approximation Model for Guaranteed QoS Model Information" – Best Student Paper Award - IEEE Symposium on Computers and Communications (ISCC'07) – Portugal, July 2007
7. W. Burakowski, M. Diaz, O. Dugeon, A. Pietrabissa, F. Racaru, G. Santoro and H. Tarasiuk. "On Multi-Domain Connection Admission Control in the EuQoS system" –

- IST mobile summit 2006 – Greece, July 2006
8. O. Dugeon et al. “*End to End Quality of Service over Heterogeneous Networks*” – In Proc. Networking Control (NetCon’05) – Lannion, November 2005
 9. O. Dugeon and A. Diaconescu. “*From SLA to SLS up to QoS control : The CADENUS framework*” – In Proc. World Telecommunication Congress / International Switching Symposium (WTC’2002 / ISS’2002) – Paris, September 2002
 10. F. Guillemin, O. Dugeon, C. Mangin, H.M. Zhang-Li, M. Basham and N. Dangy-Caye. “*Enrichment of MPLS over ATM for guaranteeing quality of service*” - In Proc. International Switching Symposium (ISS’2000) - Birmingham, May 2000
 11. F. Guillemin, J. Boyer, O. Dugeon and C. Mangin. “*Analysis of cell spacing on a TCP flow mapped onto an ATM connection*” - In Proc. Networking’2000 - Paris, May 2000
 12. F. Guillemin, J. Boyer, P. Boyer, O. Dugeon and C. Mangin. “*Cell spacing for regulating TCP over ATM*” - In Proc. International Teletraffic Congress (ITC’16) - Edhimborg, June 1999
 13. O. Dugeon. “*The RBM Simulation Machine*” – 2nd RACE Workshop – Aveiro, January 1991

4.4. Brevets

1. O. Dugeon. « *Procédé de traitement d’une requête, passerelle d’accès et système de contrôle d’admission à un service* » – INPI n° 08 54383 – Juin 2008
2. W. Htira et O. Dugeon. « *Brevet sur l’admission d’appel basé sur l’aggrégation de topologie* » - INPI n° 07 07043 – Octobre 2007
3. O. Dugeon, F. Guillemin et C. Mangin. « *Contrôle d’admission des flots TCP dans un réseau ATM par signalisation mono cellule* » - INPI n° 99 14722 – Novembre 1999
4. P. Boyer, O. Dugeon et M. Servel. « *Systèmes d’allocation d’intervalle de temps et multiplexeurs pourvus d’un de ces systèmes d’allocation d’intervalle de temps* » - INPI n° 93 09645 – Novembre 1993

4.5. Contributions aux standards

1. O. Dugeon and D. Becam, “*RACF – Requirements for reference point Rd*”, submitted and approved to RACF standard release 2 during ITU-T Q4/13 NGN GSI Reporter Group Meeting, Geneva, September 2007
2. O. Dugeon and D. Becam, “*RACF – Requirements for reference point Rr*”, submitted and approved to RACF standard release 2 during ITU-T Q4/13 NGN GSI Reporter Group Meeting, Geneva, September 2007
3. D. Becam and O. Dugeon, “*RACF – New type of CPE*”, submitted and approved to RACF standard release 2 during ITU-T Q4/13 NGN GSI Reporter Group Meeting, Geneva, September 2007
4. O. Dugeon. “*CAC requirements for the Home Gateway*”, submitted and approved for the HGI QoS release 2 during HGI plenary meeting, Lisbon, May 2007
5. O. Dugeon et al. “*CAC specification for the Home Gateway*”, presented and adopted for Home Gateway Initiative (HGI) release 2 during HGI plenary meeting, San Francisco, February 2006
6. O. Dugeon and D. Becam. “*IPTV Multicast and Unicast in RACS*” presented and

approved for RACS standard release 2 during ETSI-TSIPAN 11ter RACS R2 session, Sophia-Antipolis, October 2006

7. O. Dugeon and D. Becam “*IP TV Multicast and Unicast in RACF*” presented and approved for RACF standard release 2 during ITU-T Q4/13 NGN GSI Reporter Group Meeting, October – November 2006
8. O. Dugeon “*CAC for Home Gateway*”, presented and approved for HGI QoS release 2 during Home Gateway Initiative plenary meeting, Berlin, September 2006

4.6. Rapports de projets européens

1. O. Dugeon *et al.* “*Prototype P#3 tests report*” – IST/EuQoS D5.2.2, July 2007
2. O. Dugeon *et al.* “Home Gateway Initiative (HGI) integration in EuQoS system” in “Integrated EuQoS system Software architecture for application use cases and API” – IST/EuQoS D4.2.1 – March 2007
3. O. Dugeon *et al.* “*Tests and trial report of EuQoS system*” – IST/EuQoS D5.1.5v2, February 2007
4. O. Dugeon, E. Mingozi, G. Stea. “*Hard model Specification*” in “*EuQoS Architecture update for Phase 2*” - IST/EuQoS D1.2.2 January 2007
5. O. Dugeon *et al.* “*Demonstration trial report of EuQoS system*” – IST/EuQoS D5.1.6, December 2006
6. O. Dugeon *et al.* “*Testbed deployment report of EuQoS system*” – IST/EuQoS D5.1.4, November 2006
7. O. Dugeon. “*EuQoS architecture specification*” in “*Business models and system design analysis*” - IST/EuQoS D1.1.3, August 2005
8. O. Dugeon. “*EuQoS architecture definition*” in “*Definition of Business, Communication and QoS model*” – IST/EuQoS D1.1.1, March 2005
9. O. Dugeon and A. Brajeul. “*Project Recommendations*” – IST/CADENUS D6.3, June 2003
10. O. Dugeon *et al.* “*Service architecture with respect to eTOM processes*” – IST/CADENUS D1.3, May, 2003
11. O. Dugeon “*A preliminary cost benefit analysis*” – IST/CADENUS D6.5, February 2003
12. O. Dugeon. “*CADENUS architecture model*” in “*End-user services in the Premium IP: Models, Architectures and Provisioning Scenarios*” - IST/CADENUS D1.2, November 2001
13. O. Dugeon *et al.* “*Service Provisioning in Premium IP: Recommendation to Telecom Operators and ISPs*” – IST/CADENUS D6.2, November 2001
14. O. Dugeon. “*CADENUS system specification*” in “*Mediation components Release 1 : Requirements and architecture*” – IST/CADENUS D3.1, March, 2001
15. O. Dugeon *et al.* “*QoS control in SLA networks*” - IST/CADENUS D2.1, March 2001
16. O. Dugeon *et al.* “*SLA networks in Premium IP*” – IST/CADENUS D1.1, March 2001

4.7. Rapports internes

1. O. Dugeon. “*CAC specification for a Home Gateway*” in “*Multiservices QoS*”

- architecture for a LiveBox*” – FTR&D/RESA/GHOA, Décembre 2006
2. O. Dugeon. « *Bilan du Projet QoS TicS* » – FTR&D/CORE/M2I/O6-54, Juin 2006
 3. O. Legrand and O. Dugeon – “*Resource Control Platform Specification for IMS*” – FTR&D/CORE/NAS, Mai 2006
 4. O. Dugeon et al. “*Operax Carrier Grade Evaluation – 2nd Evaluation*” - FTR&D/CORE/M2I/62, Mai 2006
 5. O. Dugeon. « *Rapport annuel du projet QoS TicS : QoS mutualisées commandées par les services* » - FTR&D/CORE/M2I/QTC/L0.4, Décembre 2005
 6. O. Dugeon et al. “*Operax Carrier Grade Evaluation*” - FTR&D/CORE/M2I/43, Mars 2005
 7. O. Dugeon and C. Teisseire. « *Spécification du service de QoS à la demande (Qu@DS)* » – FTR&D/CORE/M2I/QTC/L2.1, Mars 2005
 8. D. Becam, O. Dugeon and A. Marrec. « *Architecture du service de QoS à la demande des services (Qu@DS)* » – FTR&D/CORE/M2I/QTC/L1.1, Janvier 2005
 9. O. Dugeon. « *Rapport annuel du projet QoS TicS : QoS mutualisées commandées par les services* » - FTR&D/CORE/M2I/QTC/L0.3, Décembre 2004
 10. O. Dugeon. « *Plan de Management de Projet (PMP) du projet QoS TicS : QoS mutualisées commandées par les services* » - FTR&D/CORE/M2I/QTC/L0.2, Mai 2004
 11. O. Dugeon. « *Dossier d’Opportunité du projet QoS TicS : QoS mutualisées commandées par les services* » - FTR&D/CORE/M2I/QTC/L0.1, Mars 2004
 12. O. Dugeon. « *Bilan du projet CER : Commande et Ressources* » - FTR&D/DAC/CPN/CER/L7.3, Février 2004
 13. O. Dugeon. « *Rapport annuel du projet CER : Commandes et Ressources* » - FTR&D/DAC/CPN/CER/L7.2, Décembre 2003.
 14. O. Dugeon. « *Rapport annuel du projet CER : Commandes et Ressources* » - FTR&D/DAC/CPN/CER/L7.1, Décembre 2002.
 15. O. Dugeon. « *Plan de Management de Projet (PMP) du projet CER : Commandes et Ressources* » - FTR&D/DAC/CPN/CER/L7.0, Décembre 2001
 16. O. Dugeon. « *Dossier d’Opportunité du projet CER : Commandes et Ressources* » - FTR&D/DAC/CPN/CER, Septembre 2001

4.8. Présentations dans des conférences internationales

1. J.F Cadiou, O. Dugeon and E. Nedellec. “Why managing QoS in Home Environment” – Broadband Word Forum Europe – Berlin, October 2007
2. O. Dugeon. “*End-to-End QoS Control over heterogeneous network*” End-to-end QoS/ QoE Workshop - ITU-T/SG 12, Geneva, 14-16 June 2006
3. O. Dugeon. “*End-to-End QoS Control*” - IEC – 21st Century Communication World-Forum - London, March 2006
4. O. Dugeon. “*EuQoS general architecture*” - Workshop EEQoS’05 – Paris, June 2005
5. O. Dugeon. « *Cadenus presentation* » – IST Mobile Cluster – Paris, 2003
6. O. Dugeon. “*A hardware Simulation Machine for Rate-Based Multiplexing Scheme Assessment: Specifications*” – RACE 1022 – November 1991

II

Architectures de gestion de la QoS

« Chaque solution amène de nouveaux problèmes »

Edward Aloysius Murphy Jr

Le modèle économique qui prévaut parmi les opérateurs de télécommunication aujourd'hui est celui du « flat rate ». Le client ne paye qu'un forfait mensuel pour un ensemble de services (VoIP, IPTV, ...) en plus de son accès à Internet. L'avenir des opérateurs de télécommunications réside donc dans leurs capacités à proposer de nouveaux services à valeurs ajoutées si ils souhaitent augmenter leurs revenus. De tels services, forcément payant en dehors du forfait, supposent un fonctionnement irréprochable donc une excellente qualité de service afin d'être adoptés par le plus grand nombre. La qualité de service est donc devenue au fil du temps un facteur clé et essentiel dans le modèle économique des opérateurs.

Cette deuxième partie décrit l'état de l'art concernant les architectures de gestion de la qualité de service, la commande de ressources et les fonctions d'admission d'appel. En particulier elle s'attache à donner les définitions des différentes notions et mécanismes manipulés. En effet, le sujet de la qualité de service est vaste, allant de la qualité perçue par l'utilisateur jusqu'à descendre finement dans la gestion des files d'attente d'un équipement de réseau en passant par la fiabilité et la sûreté de fonctionnement. Si la commande est une notion simple, le terme « ressource » couvre également une large palette de notions en télécommunication, allant de la bande passante aux adresses réseaux en passant par l'espace disque ou mémoire d'un serveur Web. L'admission d'appel intervient à différents niveaux dans la conception d'un service depuis la souscription au service jusqu'au paquet de données en passant par l'utilisation du service. Définir ces notions n'est cependant pas suffisant si elles ne sont pas reliées aux notions d'espace (partition de réseau concerné) et de temps (instant où les ressources sont configurées dans le réseau). Dans ce but, il sera tout d'abord présenté un outil de représentation des architectures afin de pouvoir classer et identifier les différents mécanismes abordés dans ce mémoire. Finalement, les différentes architectures de l'ATM, des modèles NGN et IMS, et de l'Internet sont passées en revue en tentant d'analyser leurs points forts et leur faiblesse.

1. Définitions

Avant tout, il est important de définir ce que l'on entend par **Qualité de Service** (Quality of Service : QoS) et plus particulièrement les grandeurs manipulées. De même, il est important de définir ce qu'est une ressource. Dans la suite du mémoire, on désignera par paquet l'unité de donnée minimale et élémentaire quelle que soit la technologie de transport ou la couche considérée sauf mention contraire.

1.1. La qualité de service

La qualité de service peut se définir comme étant « *l'ensemble des phénomènes pouvant influencer les performances du service qui détermine le degré de satisfaction de l'utilisateur de ce service* ». La QoS est appréhendée différemment selon l'acteur : client, opérateur de service ou de réseau et le rôle : demandeur, fournisseur, consommateur. Si elle se résume souvent à l'écoulement du trafic dans un réseau, beaucoup d'autres facteurs sont à prendre en compte. Ainsi, de nombreux facteurs comme la perte d'alimentation, la surcharge ou la panne de plate-forme de service, la rupture d'un lien de transmission, ... sont plus souvent la cause d'une mauvaise qualité perçue par l'utilisateur qu'une déficience du réseau au niveau de l'acheminement des données. Un service ou une application offrant une excellente qualité de service se doit d'être fiable, robuste et tolérant aux pannes avant même de se préoccuper du bon acheminement des données. En règle générale, un opérateur exige que la plate-forme qui rend le service affiche un MTBF (Mean Time Between Failure) supérieure ou égale à 99,999% soit une interruption de service inférieure à 5 minutes par an. Enfin, il est nécessaire de définir qui quantifie la QoS et quels sont les indicateurs mesurables à même de la vérifier. Dans le cas d'un utilisateur final qui a demandé un service, c'est la qualité perçue ou « **Quality of Experience** » (QoE) qui sera mise en avant. Dans le cas d'un opérateur de service ou de réseau, ce sont des objectifs mesurables correspondant aux paramètres de QoS qui serviront à évaluer le bon respect des engagements de qualité de service.

La gestion de la QoS, et donc sa garantie, ne peut en aucun cas se restreindre à une technologie ou une partition de réseau mais doit être abordée dans son intégralité. Elle doit former un tout et être considérée de bout en bout même si elle est ensuite subdivisée afin de rendre le problème solvable. La QoS finale est égale à la QoS rendue par l'élément le plus faible dans toute la chaîne de distribution du service. Ainsi, il ne sert à rien de garantir un niveau de QoS excellent dans le réseau de l'opérateur si la liaison Wifi du réseau domestique entre le terminal de l'utilisateur et son modem ADSL présente un taux de perte de 2%.

1.2. Paramètres de QoS

La qualité de service sera exprimée dans un langage non technique et le plus humainement compréhensible pour l'utilisateur. Par exemple, il peut être fait usage de critère de classification basé sur des niveaux (e.g. Gold, Silver, Bronze, un nombre d'étoile, ...). A l'inverse, les paramètres utilisés entre fournisseur de service et opérateur de réseau et entre opérateurs de réseaux eux-mêmes se doivent de répondre à des exigences techniques. Les paramètres de QoS [1] associés à un flux de données sont principalement :

- Le **débit** (bandwidth) du flux qui désigne la quantité d'informations écoulee par unité de temps exprimé en bit/s. Il faudra également déterminer le point de mesure auquel correspond le débit. En effet, le débit indiqué n'aura pas la même valeur suivant la couche (au sens modèle ISO) où il est mesuré, du fait des diverses encapsulations. Le débit est également associé à la notion de bande passante. Dans la suite du mémoire, la bande passante sera utilisée pour désigner la capacité d'écoulement d'un lien ou

d'un équipement et le débit pour exprimer le volume de donnée émise ou nécessaire par une application ou un service par unité de temps.

- Le **taux de perte** (loss rate et error rate) qui désigne la probabilité maximale de perte de données ou de paquets. Ce paramètre, sans unité, est bien entendu très inférieur à 1. On cherchera toujours à se rapprocher d'un taux de perte égal à 0 qui désigne une QoS excellente.
- Le **délai** (delay) de transmission, exprimé en ms, désigne le temps nécessaire pour acheminer un volume élémentaire de données de la source jusqu'à la destination. Ce paramètre peut correspondre à une valeur maximale à ne pas dépasser, une mesure moyenne ou minimale ... mais en aucun cas ne désigne le temps total de transfert des données. Il est mesuré de bout en bout ou entre deux points de référence comme étant le temps nécessaire à l'acheminement d'une unité de volume, en général un paquet, entre les deux points de mesure. Il correspond au temps de transport proprement dit (e.g. 5 μ s/km pour la vitesse de propagation de la lumière dans une fibre optique) plus au temps de traversée des différents équipements.
- La **gigue** (jitter ou delay variation) qui désigne la variation du délai de transmission, exprimée en ms.

Si le débit et le taux de perte concernent toutes les applications, le délai et la gigue affectent plus particulièrement les applications à temps réel ou requérant une grande synchronisation. Par exemple une gigue trop élevée va affecter la synchronisation de l'annulation d'échos pour un service conversationnel.

A ces paramètres, s'ajoutent des critères de disponibilité du transfert des données dans le réseau :

- La **disponibilité** du réseau qui se traduit par la probabilité qu'un élément tombe en panne. Le MTBF pour l'ensemble de la chaîne de service peut être calculé sur la base des MTBF des éléments qui la constituent,
- La **durée d'interruption** de service qui complète la disponibilité. Le MTTR « Mean Time To Repair » pour l'ensemble de la chaîne de service peut être également calculé sur la base des MTTR des éléments qui la constituent,
- La **probabilité de refus** de transfert qui est lié à la notion d'**admission d'appel**. L'opérateur peut en effet décider, en fonction du trafic présent, si son réseau a les capacités suffisantes ou non pour acheminer ce nouveau service. Cette probabilité sera maintenue la plus faible possible. À titre d'exemple, le réseau téléphonique commuté est dimensionné pour garantir un taux de rejet inférieur à 0,05 %.

1.3. Les classes de services

Ces paramètres sont ensuite regroupés entre eux en fonction des besoins des applications et des services. Ces groupes forment alors des **Classes de Services** (Class of Services : CoS). Les requêtes de QoS des applications ou des services seront toujours affectées à une classe de service donnée. A chaque classe, correspond un ensemble de paramètres de QoS avec des objectifs quantifiés. Plusieurs modèles de CoS ont été standardisés et peuvent être utilisés indifféremment. Ceux de l'ITU-T Y.1541 [1] et I.356 [2] et de l'IETF RFC 4594 [3] sont parmi les plus connus. Chaque opérateur définit également ses propres classes de services avec des objectifs quantifiés différents. Il est très complexe de synthétiser les différentes propositions. En effet, même le nombre de classes de services est sujet à des débats intenses dans les instances de standardisations. Dans la suite du mémoire, il s'agira plus de se référer au concept de classes de services qui consiste à regrouper entre-elles les applications

ayant des exigences de QoS similaires, plutôt qu'à se référer explicitement à tel ou tel modèle. Cependant, une classification des principales applications est fournie afin de faciliter la lecture de ce mémoire :

- **Voix** : Regroupe toutes les applications du type conversationnel (Voix, Visio, Conférence, ...) ayant pour contrainte forte des objectifs sur le délai et la gigue. Elles sont également sensibles au taux de perte bien qu'il ne soit pas possible de retransmettre les données et requièrent des débits assez faibles,
- **Vidéo** : Regroupe toutes les applications multimédia diffusées ou non (Vidéo à la Demande – VoD, la télévision sur IP – IP TV, ...) ayant pour contrainte forte le taux de perte et le débit et dans une moindre mesure le délai et la gigue,
- **Donnée** : Regroupe toutes les applications de transfert de données ayant pour seule contrainte un taux de perte nul et qui s'accommodent d'un délai et d'une gigue quelconque. Un débit garanti caractérise cette classe sans toutefois en faire une contrainte stricte,
- **Défaut** : Désigne toutes les applications n'exigeant aucune garantie de QoS. Bien connu sous l'anglicisme « **Best-Effort** » c'est le mode de transport du protocole IP.

1.4. Les ressources

Les ressources sont étroitement liées aux paramètres de QoS. En effet, une qualité de service donnée sera respectée si et seulement si les ressources influant sur le débit, la perte, le délai et la gigue sont bien disponibles. Si la ressource liée au débit est simple à définir¹, il n'en est pas de même pour les autres paramètres. Le délai et la gigue sont étroitement liés aux capacités d'acheminement du réseau en terme de longueur de connexion, de temps de transit dans les équipements ... Le taux de perte est lui lié aux capacités d'écoulement du trafic par le réseau.

Cependant, il existe beaucoup d'autres ressources comme celles liées aux plates-formes de service, aux terminaux ... Qu'il s'agisse de puissance de calcul, de capacité mémoire ou disque, de capacité d'affichage ou de reproduction sonore, toutes ces ressources nécessitent d'être vérifiées avant de pouvoir délivrer le service au client. Enfin, et à l'aube de l'extinction du plan d'adressage IPv4, les adresses sont également une ressource qu'il est nécessaire de gérer au fur et à mesure que le nombre de clients raccordés progresse.

A nouveau ce mémoire traitera des ressources nécessaires à la garantie de la QoS et considérera les autres comme étant disponibles et configurées.

1.5. Les contrats

Effectuer une requête de QoS ou s'engager sur un niveau de qualité impose des engagements contractuels entre le client, l'opérateur et le fournisseur de services. Un premier contrat, appelé « **Service Level Agreement (SLA)** » [4] fixe les règles d'utilisation du service par un client pour le fournisseur de services qui en retour s'engage à garantir une disponibilité de son service. Il est donc négocié entre l'utilisateur et le fournisseur de service. Afin d'acheminer correctement les données de services, le fournisseur va à son tour passer un contrat avec l'opérateur. Appelé « **Service Level Specification (SLS)** » [5], il fixe les règles d'usage des capacités du réseau par le fournisseur de service pour l'opérateur qui en retour s'engage à garantir une QoS pour l'acheminement des données du service. En général, le SLS est considéré comme étant l'annexe technique du SLA bien qu'il ne soit pas vu de l'utilisateur final étant ajouté par le fournisseur de service. La garantie de service et de QoS ne sera

1 Elle correspond à la capacité disponible sur les liens et les équipements

assurée que dans le cadre de ces contrats. Par exemple si un SLS spécifie que l'application doit envoyer un trafic à 2 Mbit/s mais que le serveur émet un flux à 2,5 Mbit/s, l'opérateur est en droit de détruire le trafic excédent afin de protéger non seulement son réseau mais également les autres contrats pour lesquels il est engagé. Le SLA utilise les paramètres non techniques de haut niveau pour décrire la QoS alors que le SLS est composé de paramètres techniques tels que décrits au §1.3.

Les opérateurs utilisent la même notion de contrat lorsqu'ils souhaitent s'échanger du trafic. Ils sont également appelés « **Peering Agreement** » ou « **Accord de Peering** » car ils sont souvent réciproques dans le sens où il s'agit d'un échange mutuel et si possible symétrique.

1.6. Flot, micro flot, agrégat et session

Un flux ou flot de données correspond à l'ensemble des paquets échangés entre une même source et destination et appartenant à la même application. Un micro flux ou micro flot est une subdivision d'un flot. Par exemple, une application de visioconférence transfère un flot composé d'un micro flot audio et un micro flot vidéo. Un flot a un début et une fin. Une application peut comporter plusieurs flots.

L'agrégat correspond à l'ensemble des paquets ayant la même destination et appartenant à la même classe de service. Les paquets peuvent appartenir à plusieurs flots.

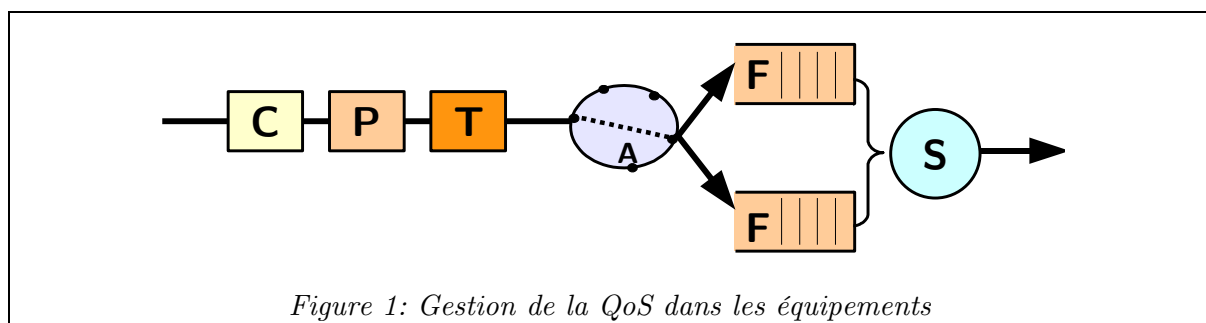
Une session correspond à l'activité d'une application comprise entre un début et une fin. La session est généralement matérialisée par le protocole applicatif e.g. SIP, RTSP, ... Elle peut être courte, transfert d'une page HTML, ou longue, visualisation d'une vidéo.

2. Fonctions élémentaires de gestion de la QoS

2.1. Décomposition des équipements

En premier lieu, la QoS doit être garantie au travers de la maîtrise des ressources dans le plan de transfert. Ce niveau regroupe l'ensemble des mécanismes inhérents aux technologies de transfert tel que l'ATM, la SDH, le GigaBit Ethernet, le GPON, ... mais également les disciplines de gestion des files d'attente des équipements. Il est symbolisé par 5 fonctions élémentaires (voir figure 1) au sens large, que l'on retrouve dans toutes les technologies dans l'ordre suivant du traitement des données :

- La **classification** (C) ou identification qui consiste à identifier le flux de données. Cette fonction utilise tous les types d'identifiants disponibles selon la technologie de transport. Ils sont soit logique (VP/VC pour l'ATM, LSP pour MPLS, flow label en IPv6), soit du type adresse (E.164, IPv4, IPv6, MAC, ...) soit du niveau marqueur (code DSCP). Il est quelquefois nécessaire de remonter dans les couches (numéro de port) afin de compléter ou distinguer au sein d'un flux différentes composantes.
- La **police** (P) qui vérifie la conformité des flux de données par rapport à un contrat de trafic.
- La **traduction** (T) ou marquage qui intervient si les données sont jugées conformes pour traduire l'identifiant logique (VP/VC, label MPLS) ou marquer le flux avec un marqueur (code DSCP).
- L'**aiguillage** (A). L'aiguillage regroupe l'acheminement du paquet vers la bonne sortie soit en fonction de l'identifiant logique éventuellement modifié par la fonction de traduction pour la commutation soit sur la base de l'adresse de destination dans le cas du routage bond par bond.
- Les **Files d'attentes** (F). Placées en entrée et/ou en sortie des équipements, elles permettent d'adapter l'arrivée des flux au rythme de sortie et temporisent l'accession au lien de sortie ou tout autre organe interne de l'équipement comme la fonction d'aiguillage. Elles sont systématiquement associées avec une discipline de service d'entrée et de sortie.
- Finalement le **serveur** (S) gérant la ou les files d'attentes décide du paquet de données à expédier sur le lien de sortie. Les disciplines de services sont diverses et variées : priorité stricte, priorité, round robin, waited fair queing, espacement ... selon qu'elles prennent en compte une ou plusieurs files d'attente correspondant à une classe de service donnée.



Toutes les fonctions ne sont pas forcément présentes dans un équipement. Par exemple, il peut ne pas supporter les fonctions de classification et police et implémenter qu'une seule

discipline de service assurant ainsi que la fonction de commutation ou de routage simple. Le support de la QoS est dans ce cas là quasi inexistant. De même, l'ordre des fonctions illustré sur la figure n'est donné qu'à titre d'illustration. Une première couche de file d'attente peut résider en entrée avant tout traitement, avant l'accès à la fonction d'aiguillage,.... Cependant, certaines fonctions élémentaires sont indissociables. La classification/identification interviendra toujours avant la police car il faut pouvoir disposer du contexte inhérent au flux de données traité pour vérifier sa conformité. Une file d'attente est toujours associée à une discipline de service. L'aiguillage ou le routage nécessite obligatoirement un marquage ou une traduction. Notons que la fonction de classification peut également porter sur toutes les informations couvrant les niveaux 1 à 7 du modèle ISO. Il est ainsi possible de repérer un flux par type d'application (comme le peer-to-peer) même si il n'utilise pas un mode de transport classique (au sens numéro de port et protocole habituel). Ce modèle est connu sous le nom de « Deep Packet Inspection (DPI) » par opposition aux méthodes traditionnelles de classification que l'on retrouve dans la plupart des équipements de réseau. Les files d'attentes peuvent être de grande taille. Elles conviennent bien aux applications exigeant un taux de perte nulle. Par contre, le temps d'attente augmente significativement ce qui ne convient pas aux applications exigeant un délai et une gigue faible. A l'inverse, les files d'attentes courtes garantissent un temps d'attente faible favorisant le délai et la gigue, mais au détriment d'un taux de perte plus élevé.

Les flux de données vont donc subir, au cours de la traversée de ces 5 fonctions, des traitements à même d'influer positivement ou négativement sur la QoS.

2.2. Les différentes méthodes de gestion de la QoS

2.2.1. Surdimensionnement des réseaux

En premier lieu, une technique simple pour garantir de la QoS est de surdimensionner son réseau. Dans ce cas, la bande passante peut être considérée comme quasiment illimitée. Dans la réalité, un réseau dit surdimensionné offre environ une capacité 10 fois supérieure au trafic qu'il doit écouler. Cette surcapacité permet également de garantir un taux de perte quasi nul puisque le réseau est à même d'acheminer tout type de trafic. Restent le délai et la gigue qui si ils sont faibles n'en demeurent pas moins non maîtrisés ni garantis. Cela s'avère crucial lorsque la distance est grande, tant en terme de domaines et équipements traversés qu'en distance géographique. Il est illusoire de croire que le délai et la gigue peuvent être garantis sur des réseaux surdimensionnés dès lors que le nombre de réseaux et le temps de transit de ces réseaux ne sont pas bornés. Les applications appartenant à la classe de service Voix seront directement impactées par un tel comportement. En résumé, seule la classe de service « Best-Effort » est supportée. C'est le modèle de l'Internet tel que nous le connaissons.

2.2.2. Modélisation et multiplexage statistique

La deuxième méthode la plus en vogue et la plus utilisée consiste à modéliser la superposition des flux afin de déterminer le dimensionnement des files d'attentes et des paramètres de police à configurer dans les équipements. Connue sous le vocable de multiplexage statistique, cette méthode est basée sur le constat qu'il est possible de déterminer par un modèle mathématique l'enveloppe de trafic résultant de la superposition de n flux. Très en vogue dans les années 90, elle a donné lieu à des prouesses mathématiques et des recherches théoriques très poussées dans le domaine de la gestion et la modélisation des files d'attentes. Cependant, si il est possible de modéliser la superposition de flux de donnée suivant des lois simples (comme la loi de Poisson), cela devient très complexe, voir impossible

lorsqu'il s'agit de lois complexes ou cherchant à s'approcher le plus possible du comportement réel des applications. Le multiplexage de flux suivant des lois différentes rend le problème encore plus complexe à résoudre de même que l'application des modèles à l'échelle d'un réseau. Certes, il est nécessaire de pouvoir dimensionner un réseau et un équipement, mais cela doit rester dans des proportions raisonnables en terme d'implémentation. Il n'est pas concevable d'exécuter un algorithme complexe pendant plusieurs minutes pour déterminer si un flux peut être acheminé ou non dans le réseau.

2.2.3. Commande des équipements

La dernière méthode qui correspond au cœur de ce mémoire, concerne la commande dynamique des différentes entités élémentaires constituant un équipement en fonction des besoins du flux. L'application nécessitant de la QoS se voit ainsi offrir un traitement adapté à ses besoins. La configuration est soit **semi statique** lorsqu'il s'agit de traiter les flux par classes de services. Seule la classification est éventuellement configurée de sorte à marquer les flux pour qu'ils soient aiguillés vers la bonne file d'attente. Soit **dynamique** lorsque les fonctions de classification, de police et de marquage sont configurées spécifiquement pour ce flux.

La gestion de la QoS peut ensuite être regroupée selon deux grandes écoles : soit par **session**, soit par **classe de service**. Les paramètres de QoS seront alors appliqués soit aux flots ou micro flots appartenants à la session, soit à l'agrégat correspondant à la classe de service. Dans le premier cas, il est possible de garantir une QoS **absolue** et individualisée alors que dans le deuxième mode il ne peut s'agir que d'une garantie **relative** par rapport aux autres classes de services. Il n'est fait aucune distinction entre les flots ou micro flots au sein d'une même classe. Il est généralement nécessaire de devoir maintenir connus les états de chaque session alors qu'il n'est pas nécessaire de connaître le détail des flots dans la gestion par classe. C'est donc bien deux philosophies opposées qui prétendent gérer la QoS dans les réseaux. D'un côté les architectures basées **session** et de l'autre les architectures basées **classe de service**.

2.3. Les points de congestions

Ce sont avant tout les pannes et les saturations des plates-formes de services qui sont les premières causes de dégradation de la qualité de service. Puis, au sein des réseaux, c'est la perte de paquet ou le non respect du délai et de la gigue. Ces dysfonctionnements sont dus à ce qui est communément appelé **point de congestion**. Il s'agit d'un endroit précis dans le réseau où l'écoulement du trafic est perturbé et où les garanties de QoS ne sont plus respectées. Ils sont localisés soit dans les équipements soit sur les liens de raccordements, bien qu'il soit plus approprié de parler de l'interface de sortie de l'équipement que du support physique du lien de raccordement. Cependant, dans le cas de transmission radio ou par porteuse sur ligne électrique, les perturbations électromagnétiques peuvent modifier les caractéristiques du lien de transmission et donc engendrer des pertes de données.

Au sein des équipements, les pertes sont localisées soit dans les files d'attentes (F) soit dans la fonction d'aiguillage (A) lorsque la matrice de commutation ou de routage est de type bloquante. C'est-à-dire que la capacité totale de commutation ou de routage de la fonction d'aiguillage exprimée en débit est inférieure à la somme des débits entrant sur toutes les interfaces de l'équipement. Dans tous les cas, l'aiguillage ou la file d'attente se voit présenter une charge $\rho > 1$. Si la période de surcharge est faible, la ou les files d'attentes sont en mesure d'absorber la pointe de trafic. Dans le cas contraire, on observe un débordement des mémoires de stockages et donc une perte de données. Cette surcharge engendre

systématiquement un délai d'attente supplémentaire qui impacte directement le budget total du délai de traversée du réseau et, en cas de surcharge à répétition, une gigue de multiplexage élevée. À noter que en cas de pertes par débordement, tous les flux de données présents dans la file d'attente ou l'aiguilleur sont potentiellement touchés et pas seulement les flux fautifs ou en excès.

Une dernière source de perturbation vient de la mauvaise configuration des équipements ou du non respect du contrat de service par l'application. Dans les deux cas, la fonction de police (P) va soit marquer soit détruire des paquets jugés non conformes. La fonction de classification (C) peut également ne pas reconnaître correctement les flux et donc déclencher un mauvais contrôle par le police ou un mauvais marquage. Finalement, si les données ne sont pas marquées (fonction M) avec le bon marqueur, elles ne seront pas aiguillées dans la bonne file d'attente et subiront un service inapproprié par rapport à la QoS demandée.

2.4. Principes de base à respecter

En résumé, la QoS sera assurée au niveau d'un équipement si :

- Le débit nécessaire à l'acheminement du flux de données est disponible sur le lien de sortie et dans l'équipement que ce soit au niveau de la fonction d'aiguillage ou dans les files d'attentes,
- Le contrat de trafic est respecté par l'application et que les paramètres correspondants pour la fonction de police aient été correctement ajustés,
- Le temps de transit dans l'équipement et sur le lien de raccordement est le plus faible et le plus stable possible,
- Le flux de données est aiguillé dans la bonne Classe de Services pour être traité selon la priorité requise par l'application.

Les deux premiers points garantissent un taux de perte minimal et un débit donné alors que les deux derniers garantissent un délai et une gigue minimale. Ces contraintes doivent ensuite être appliquées à l'ensemble des équipements et des réseaux.

Enfin, il est souhaitable de pouvoir isoler au mieux les flux entre eux afin d'éviter qu'un flux fautif ou en excès engendre des pertes sur tous les autres flux. De même, il est nécessaire de pouvoir contrôler les flux admis dans le réseau si la garantie de QoS n'est pas réalisable pour ce flux non seulement pour protéger le réseau mais également pour protéger les flux déjà en cours d'acheminement.

3. Outil de représentation des architectures

3.1. Le modèle OURAf

Au fur et à mesure que les études et les travaux de recherches progressaient et inventaient de nouvelles solutions, les opérateurs de télécommunications se sont retrouvés devant une jungle de solutions toutes plus compliquées les unes que les autres. Afin de trier efficacement les protagonistes et discerner les bonnes des mauvaises solutions, ou du moins les mieux adaptées à un contexte donné, il est devenu indispensable de disposer de critères de classifications. De façon analogue à l'urbanisme urbain, les opérateurs ont commencé à étudier l'urbanisme des réseaux vers la fin des années 90 dans le but de fixer des règles de conception d'architecture. C'est dans ce contexte que le modèle OURAf (Outil de Représentation d'Architecture fonctionnelle) [6], [7] a été mis au point dans le but de cartographier les architectures de réseaux. D'autres modèles d'urbanisme ont été élaborés, comme celui du TMF [8] ou de l'ITU [9] mais ils sont en général moins complets et plus adaptés au système d'informations et la gestion des réseaux. Cet urbanisme des réseaux va donc plus loin que la thématique de la qualité de service. Aussi, après une rapide présentation, seule la partie propre à la gestion de la QoS sera détaillée. Il est à noter qu'un modèle de représentation doit être neutre vis à vis des choix architecturaux.

3.2. Principe

Le modèle est basé sur la relation client serveur. Un **client** effectue une demande, une requête, auprès du **serveur** concerné, pour obtenir un **service**. Cette relation met en jeu des fonctions, des données associées, et des échanges de données entre le client et le serveur, ainsi qu'au sein des domaines client et serveur. Ces échanges sont identifiés par des points de référence dans une représentation purement fonctionnelle, et par des interfaces dans une implémentation organique. Le client peut être un client final ou un serveur en position de client, ce qui assure la récursivité du modèle. On distingue également le client de l'utilisateur dans le sens où le client est la personne qui est facturée alors que l'utilisateur est le consommateur du service. Si dans la majorité des cas il s'agit de la même personne, cette notion est importante dans le modèle de données sous-jacent. En effet, plusieurs utilisateurs peuvent utiliser un service souscrit et donc payé par un seul client. Il est donc nécessaire de pouvoir distinguer l'ensemble des utilisateurs. Un très bon exemple correspond au client qui s'abonne au service d'accès à l'Internet sur ADSL pour l'ensemble des membres de sa famille qui forment les utilisateurs finaux.

Il s'agit d'inventorier et de classer logiquement, du général vers le particulier, l'ensemble des fonctions, des données associées, des points de référence, intervenant dans les réseaux de télécommunications (transport, commande et gestion).

3.3. Les niveaux

La découpe en niveaux est obtenue compte tenu d'une certaine chronologie des traitements, de la requête du client jusqu'à la mise en œuvre des ressources, avec le souci de séparer les opérations relatives : au client, au service, aux traitements des ressources et aux ressources proprement dites (voir figure 2). Ces niveaux sont regroupés selon :

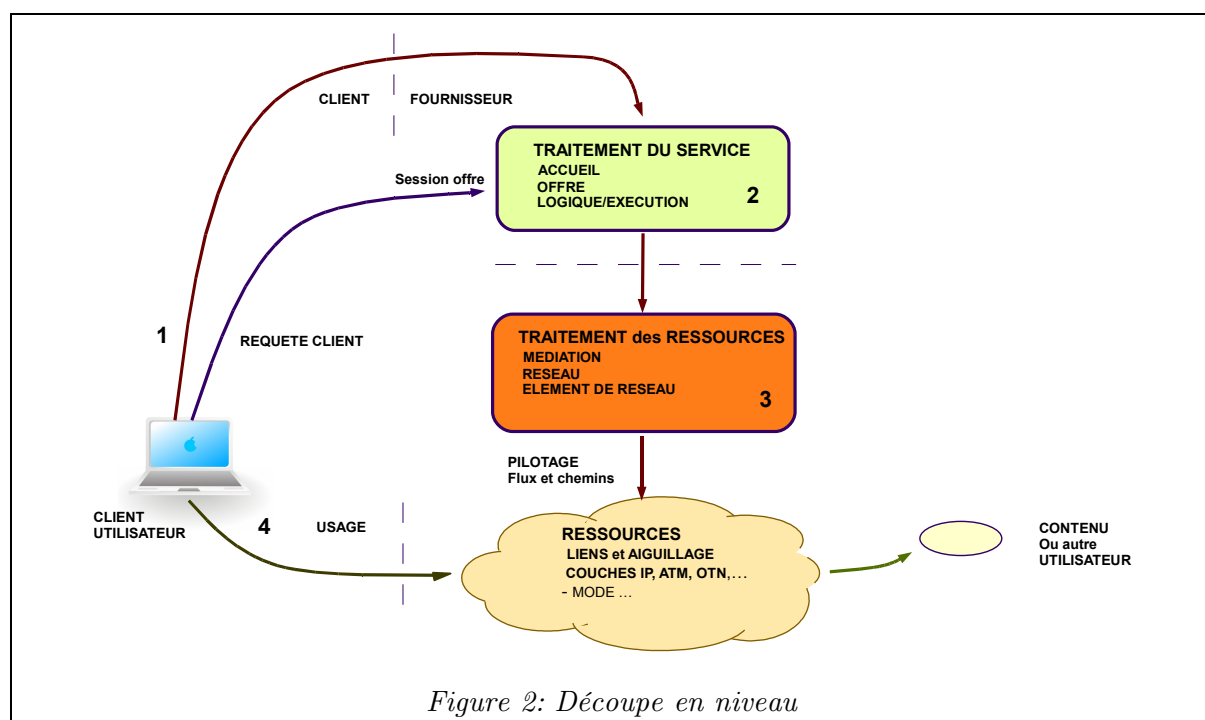
- le **traitement des services** (et du client) qui comprend les niveaux : accueil, médiation de service, logiques de service,
- le **traitement des ressources** qui comprend les niveaux : médiation d'accès aux ressources, traitement des ressources réseau, traitement des ressources élément de

réseau,

- les **ressources** proprement dites de transport et de contenu.

3.3.1. Traitement des services

Il regroupe 3 niveaux qui gèrent l'ensemble de la relation entre le client et le service. Tout d'abord, le niveau **accueil** appelé « **Médiation d'Accès au Service** » ou « Front Office » est l'intermédiaire entre le client au sens large et les offres de service. C'est le niveau qui gère la première page d'accueil du client lui assurant l'accès à l'ensemble des fournisseurs d'offres. Ensuite, le niveau **offre** appelé « **Médiation de Service** » ou « Back Office », agrège les offres de services sous une forme unifiée et assure la présentation du catalogue des services, les souscriptions et la mise en relation de l'utilisateur avec la logique du service concerné. Finalement, le niveau **service** est appelé également logique de service, service unitaire, instance de service ou invocation. Il permet de traiter les données d'entrée relatives à un service requis (identifiants, description du trafic, qualité,...), d'exécuter la logique du service requis et délivrer les données de niveau applicatif (Voix, VoD, jeux, consultation,...).



3.3.2. Traitement des ressources

De façon similaire au traitement des services, il regroupe 3 niveaux qui gèrent le plan de contrôle du réseau. Le premier niveau appelé **médiation d'accès aux ressources** sert d'intermédiaire entre les logiques de services et les ressources support. Il fait la traduction service – ressources à partir des données caractéristiques du service requis (identifiants, débit, QoS, ...), et des ressources de bout en bout (terminaux et ensemble des sous réseaux support du service), nécessaires à la mise en œuvre du service. A cette fin il dispose des données décrivant ces différentes ressources (localisation, capacités en matière de connectivité, de QoS, ...). Il transmet aux ressources identifiées les contraintes de mise en œuvre, en conformité avec les contraintes du service. La médiation d'accès aux ressources est indispensable du fait : de l'indépendance service ressource, du morcellement des réseaux, des collisions de plans d'adressage, des offres de QoS variées, ...

Elle va ensuite piloter le niveau appelé **ressources réseau** qui permet de planifier et de déployer les ressources des sous réseaux, de configurer les ressources des sous réseaux, et de fixer les conditions d'acheminement et d'écoulement des flux d'information, compte tenu des données prévisionnelles de trafic et des mesures de charge. Elle identifie le chemin topologique et le type de contrôle de flux à appliquer suite à une demande relative à un service et effectue les contrôles d'admission en comparant les contraintes de la demande aux disponibilités du sous réseau. Elle collecte les données de performance et d'usage en vue de l'élaboration de la QoS et de la facturation. A noter que l'acheminement réseau, ou topologique, consiste à identifier les nœuds et les liens (interface) à partir d'un adressage global (IP, E164).

Enfin, le dernier niveau s'intéresse aux **ressources des éléments** de réseau, c'est-à-dire aux équipements le constituant. C'est la déclinaison par technologie et par équipements du niveau précédent. De plus, il positionne les ressources de transport proprement dites : terminaisons de liens et fonction aiguillage. A noter que l'acheminement dans l'élément de réseau, consiste ici à identifier, à l'intérieur du chemin topologique déterminé ci-dessus, et dans chaque nœud d'aiguillage, le chemin virtuel à partir d'un adressage local (MPLS, VP, VC, VLAN, ...).

3.3.3. Ressources

Dernier niveau dans le modèle, il s'agit ici des **ressources de transport** et de contenu. Les ressources de transport concernent les capacités d'aiguillage des nœuds, les capacités d'écoulement des flux sur les liens, prises en compte à travers les notions de débit ou de bande passante, et gérées dans des files d'attente permettant le contrôle de ces flux, et les capacités de filtrage, police et marquage des flux. Les ressources de contenus concernent les capacités de stockage des serveurs et les adaptations de codage des flux multimédias.

3.4. Les Processus

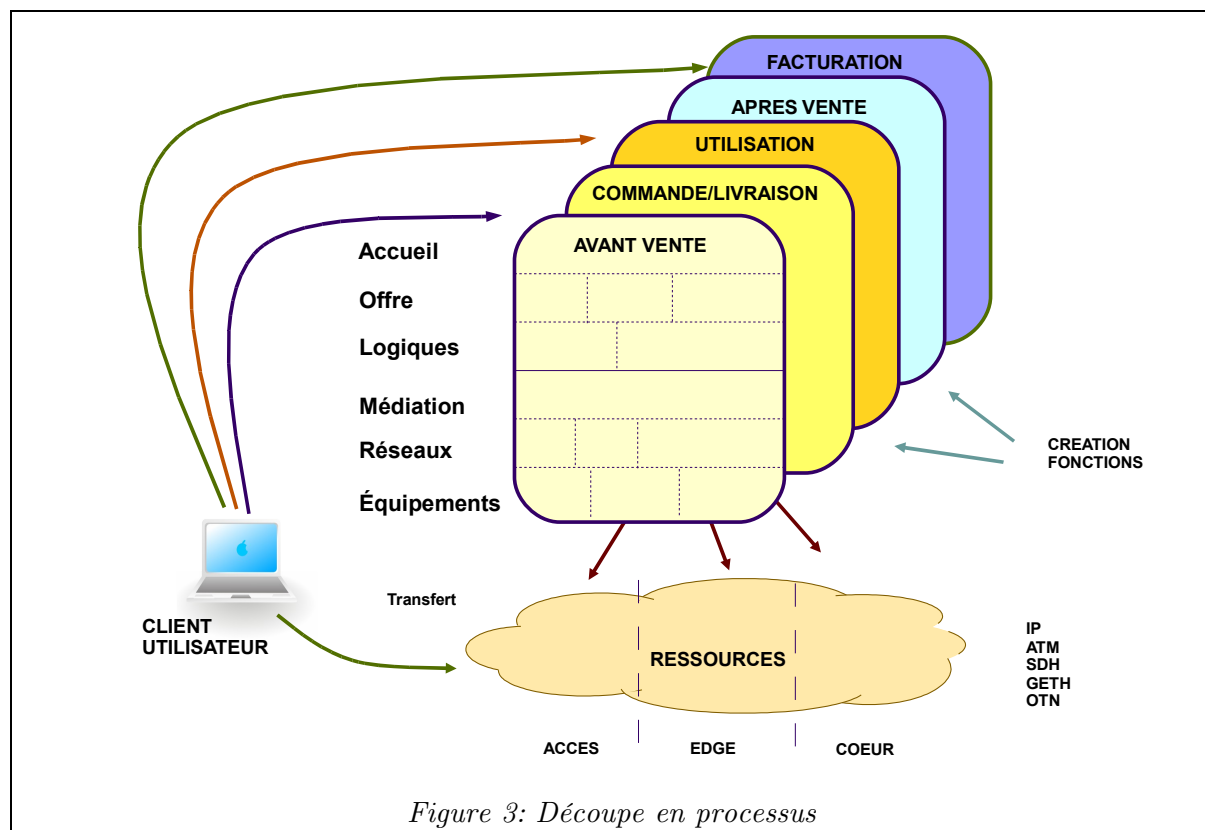
Les processus, (voir figure 3), s'inscrivent également dans une chronologie. Ils sont issus d'une structuration des actions entreprises dans les réseaux suite aux différentes opérations demandées par le client, de l'avant vente à la facturation du service, en passant par la commande – livraison, l'utilisation et l'après vente. Chaque processus est découpé selon les niveaux décrits précédemment. Les processus décrits ici sont des processus orientés client final. Cependant la découpe reste valide quand l'activateur est un fournisseur quelconque en position de demandeur de service, ou un opérateur en situation d'exploitation – maintenance de son réseau. Le modèle définit les notions de processus client, de processus fournisseur, et de processus opérateur. Par ailleurs, la notion de processus doit être complétée par les aspects création de service, et plus généralement par la création hors ligne de toute fonction identifiée dans les niveaux et processus, et qu'il s'agit d'introduire ou de modifier en ligne.

Cinq processus sont identifiés :

- **l'avant vente**, soit l'ensemble des opérations, préalables à la décision de souscription de service par les clients,
- la **commande / livraison**, soit l'ensemble des opérations liées à la souscription des services et à la configuration des ressources en conséquence, appelée également approvisionnement ou « provisioning » en anglais,
- **l'invocation** ou **l'utilisation**, soit l'ensemble des opérations liées à une demande d'usage d'un service,
- **l'après vente**, soit l'ensemble des opérations liées au traitement des données d'usage

d'un service et à la supervision du comportement du réseau, appelée souvent maintenance ou SAV,

- la **facturation**, soit l'ensemble des opérations liées aux traitements des données caractéristiques de la consommation, par le client, des ressources du réseau.



Une des innovations du modèle OURAf est la création du processus d'invocation ou d'usage. En effet, si les autres modèles, plutôt orientés système d'information de l'opérateur, définissent bien les 4 autres processus, ils font l'impasse sur l'utilisation du service. C'est pourtant durant ce processus que se déclenche les signalisations de services (e.g. SIP, HTTP, ...) et, dans le cadre de ce mémoire, la commande des ressources associées aux transferts des données. La distinction entre le client et l'utilisateur trouve toute sa signification avec la découpe en processus. En effet, c'est bien le client seul qui sera sollicité lors des processus d'avant-vente et de facturation et uniquement l'utilisateur lors du processus d'invocation. Les processus de commande / livraison et d'après-vente pouvant concerner aussi bien le client que l'utilisateur.

La gestion de la QoS, dans l'outil OURAf, est concernée par les seuls processus de commande/livraison pour l'approvisionnement des ressources, l'invocation pour la commande des ressources et l'après-vente pour la surveillance et la mesure de la QoS. Le processus de **commande / livraison** rassemble, dans les différents niveaux, l'ensemble des actions menées suite à la décision d'un client de souscrire à un service. Ces actions concernent l'accueil du client souscripteur, l'ensemble des opérations liées à la souscription des services (contrats, profils), la gestion des services souscrits, et la gestion des configurations des ressources de bout en bout, de sous réseau et d'élément de réseau.

Le processus d'**invocation** rassemble, dans les différents niveaux, l'ensemble des actions menées suite à la décision d'un utilisateur de faire usage d'un service nécessitant l'utilisation des ressources de transport. Ces actions concernent l'accueil du client utilisateur de services, les contrôles d'admission par rapport au profil souscrit, le traitement des données liées à l'utilisation, le traitement des ressources impliquées (de bout en bout, de sous réseau et d'élément de réseau) et l'usage des ressources pour le transport des flux data,... Dans ce processus, le traitement des services et des ressources se nomme respectivement **commande de service** et **commande de ressources**.

Le processus d'**après-vente** rassemble, dans les différents niveaux, l'ensemble des actions menées, en particulier, suite à l'activation du processus d'utilisation. L'utilisation génère des données d'usage qu'il s'agit de traiter en agrégeant les résultats : de QoS, de statistiques relatives à un client, de mesures de performances, de collecte des anomalies et des fautes (de bout en bout, de sous réseau et d'élément de réseau) afin de mener à bien des actions de supervision, de mesure et de maintenance.

3.5. Les partitions

Les **partitions** correspondent à une découpe verticale de chacun des niveaux, dans chacun des processus. Il y a indépendance entre la découpe en partition d'un niveau, par rapport à la découpe en partition d'un autre niveau. Les partitions dans le niveau ressources correspondent, par exemple, aux notions d'accès, de périphérie de service, de cœur de réseau, ... Les partitions les plus significatives concernent les ressources des réseaux : locaux, d'accès, de collectes, de cœurs ... Mais les autres niveaux sont également segmentés, fournisseurs de bout en bout et fournisseurs de services par exemple, pour constituer la « chaîne de valeurs » support d'un service donné. Au sein même de ces partitions, les points d'interconnexions avec les autres partitions seront l'objet d'une attention toute particulière. Ils délimitent également la zone où s'exerce la responsabilité juridique de l'opérateur. Par la suite, cette zone sera appelée **domaine** ou AS (Autonomous System). Un opérateur pouvant contrôler et gérer plusieurs domaines. De même un client peut déléguer la gestion de son domaine à un opérateur : c'est le cas des clients entreprises lorsqu'ils demandent une prestation d'infogérance. Les points d'interconnexions entre partitions sont appelés intra domaines lorsqu'ils sont à l'intérieur d'un même domaine et inter domaines lorsqu'ils matérialisent la frontière entre deux opérateurs e.g. entre deux réseaux cœurs, entre un réseau local et le réseau d'accès, i.e. dès qu'il y a changement de juridiction. Ainsi, en fonction de la partition, les mécanismes de gestion de la QoS seront différents ou tout du moins configurés différemment.

3.6. Représentation fonctionnelle

La découpe fonctionnelle (voir figure 4) en sept niveaux et en cinq processus sous forme de matrice obéit à des contraintes : de chronologie des traitements suite à une demande du client, de séparation des opérations liées au client, au service, au traitement des ressources, et aux ressources proprement dites. Il en résulte un espace structuré décrivant progressivement le monde des télécommunications : en niveaux ou en processus, en unités fonctionnelles, en fonctions élémentaires et selon les données associées et les règles d'échange entre les différentes unités.

Cette grille de référence fonctionnelle, à mailles variables, i.e. à plusieurs degrés de granularité des descriptions, sert à la description du contenu de différentes entités : rôle, acteur, équipement, projet,... Elle est neutre en ce sens qu'elle ne préjuge pas d'un choix d'architecture cible donné.

4. La commande des ressources

4.1. Application de OURAf

OURAf a permis, entre autre, de définir précisément en normalisation [10] les différents processus et niveaux composant l'architecture de la gestion de la QoS pour les réseaux NGN et en particulier dans le cadre de l'IMS mais reste valable pour les autres architectures de réseau. La figure 5 illustre les différentes étapes dans l'approvisionnement, l'utilisation et la surveillance de la commande de la QoS à l'aide de l'outil OURAf.

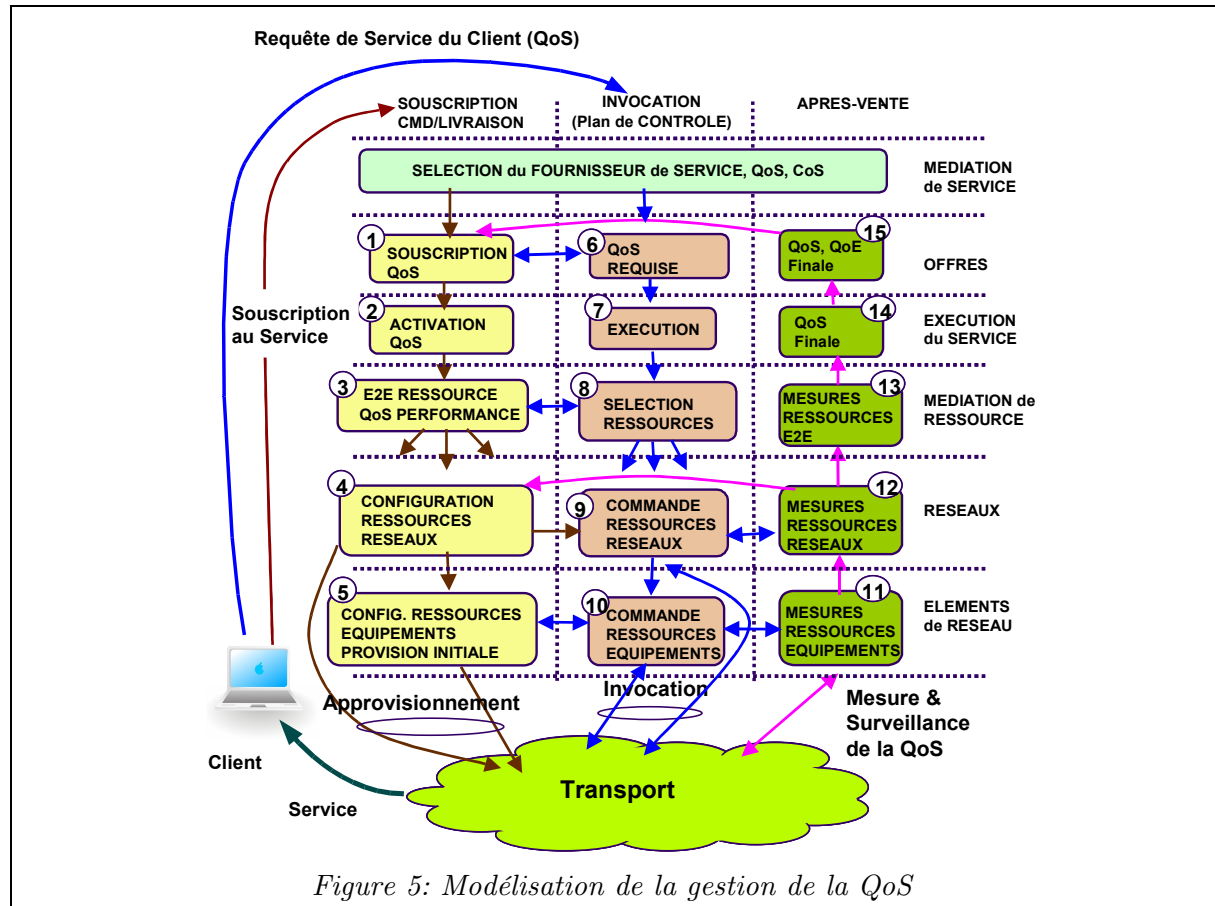


Figure 5: Modélisation de la gestion de la QoS

La demande de l'utilisateur se décompose donc en 3 temps en fonction du processus utilisé. Tout d'abord, **étapes 1 à 5**, l'utilisateur va souscrire au service. Cette souscription va déclencher durant le processus de commande / livraison la mise en place des ressources réseaux et la configuration des équipements une fois la vérification de la faisabilité de ce nouvel abonnement. Par exemple, l'opérateur va vérifier si la ligne xDSL du client est à même de supporter, en terme de débit, l'abonnement au service TV. La phase d'approvisionnement inclut également la création des profils de l'utilisateur dans les différentes bases de données gérant le service et le client. Ce sont les niveaux du traitement du service qui sont sollicités pour cette première phase. Les niveaux du traitement de ressources entre en jeux lors de l'approvisionnement des ressources proprement dit.

Une fois l'abonnement enregistré et configuré, le client peut commencer à utiliser son

service représenté par les **étapes 6 à 10**. La première étape consiste à accueillir le client et lui proposer les différentes offres liées à son abonnement. Cela peut aller de la liste des fournisseurs de contenus Vidéo ou TV qu'il peut consulter à une simple tonalité d'invitation à numéroter. Une fois le contenu ou service sélectionné, la dernière étape du traitement de service consiste à solliciter le traitement des ressources afin de i) vérifier la capacité du ou des réseaux vis-à-vis des ressources demandées, ii) les réserver et iii) les configurer.

Durant la consommation du service, **étapes 11 à 15**, le processus d'après-vente vérifie que la QoS délivrée est conforme à celle demandée. Il s'agit essentiellement de corréler les mesures effectuées sur les équipements, dans les réseaux et de bout en bout avec la qualité réellement perçue par l'utilisateur. Appelée « **Quality of Experience - QoE** », elle est plus connue dans le monde de la téléphonie sous le vocable de note MOS (Mean Opinion Score) [11]. En cas de problème, les différents éléments de la chaîne sont capables d'interagir afin de rétablir le niveau de QoS.

4.2. Approvisionnement

La première étape à laquelle un opérateur est confronté, est l'approvisionnement de son réseau. L'ensemble des actions liées à cette mise en place des ressources se situe au sein du processus commande / livraison et plus particulièrement dans les niveaux traitement de ressources et ressources du plan de transfert. Cet approvisionnement des ressources peut résulter d'un nouvel abonnement ou souscription, mais il est également déclenché par l'opérateur lui même lors des phases de construction et de planification du déploiement de ses réseaux. Processus de gestion, la commande / livraison est directement traitée par le système d'information (SI) de l'opérateur. Il regroupe dans des bases de données aussi bien la connaissance des réseaux, les profils des clients et utilisateurs que les méthodes à appliquer pour la construction et l'approvisionnement des réseaux.

Pour bien comprendre la différence avec le processus d'invocation, il est nécessaire de faire intervenir le facteur de l'échelle de temps à laquelle se déroule la commande des ressources. Cette phase peut prendre de quelques minutes à plusieurs jours voire semaines sans que les performances vis-à-vis de l'utilisateur soient dégradées. L'échelle de temps dépend étroitement du type de construction qu'il faut réaliser dans le réseau. Il est facilement compréhensible que le temps nécessaire à l'ajout d'un équipement ou d'une liaison soit plus long que le simple ajout d'une configuration dans un équipement. Il n'en demeure pas moins que ce processus peut être relativement lent par rapport aux actions à entreprendre dans le processus d'invocation où l'utilisateur souhaite consommer son service immédiatement. Le but principal est de pré réserver et configurer des ressources dans les équipements qui seront utilisées, consommées, ultérieurement par les services invoqués par les utilisateurs lors du processus d'invocation. Cette configuration des équipements s'effectue par la gestion depuis les plates-formes du système d'information de l'opérateur en utilisant des protocoles comme SNMP [12] ou plus récemment en XML [13].

La deuxième méthode la plus utilisée est l'ingénierie trafic. Les modèles mathématiques et la planification sont alors utilisés pour dimensionner le réseau, les plates-formes de services, la capacité des liens, ... en fonction du nombre d'utilisateurs. Les règles d'ingénieries produites peuvent être :

- **Statiques** : Elles indiquent à la fois quand il faut ajouter des équipements, liens, plates-formes, et leurs configurations.
- **Dynamiques** : Sur un réseau donné, elles servent à la construction des connexions semi statiques du type VP ATM ou tunnels MPLS.

4.3. Invocation

4.3.1. Classification de la commande de ressources

La commande des ressources au sein du processus d'invocation se décline dans beaucoup d'architecture. Il est donc nécessaire à l'aide de l'outil OURAf de pouvoir classer les différents modes de commande par famille. Dans un premier temps, il faut identifier les acteurs et leurs relations. La figure 6 illustre les 3 niveaux du traitement du service et les 3 niveaux du traitement des ressources du modèle OURAf ainsi que le plan de transfert matérialisé par les équipements. Deux partitions de réseau sont également représentées pour matérialiser, dans le plan de transfert, à la fois l'interface usage/réseau ou « User Network Interface » (UNI) et l'interface réseau/réseau ou « Network Network Interface » (NNI). L'interface UNI existe également dans le niveau service et ressource.

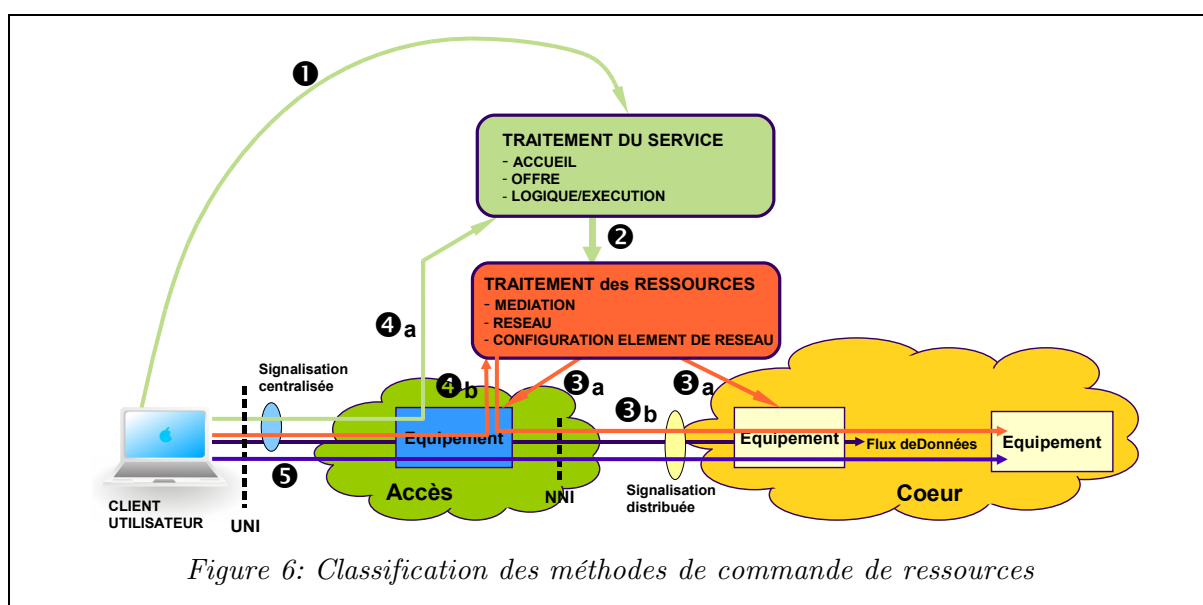


Figure 6: Classification des méthodes de commande de ressources

Soit les interactions suivantes :

1. **Demande de service de l'utilisateur.** Les paramètres de QoS sont explicites (sous forme de niveau) ou implicites (déduits du service). La requête n'est pas portée par une signalisation usage / réseau à l'UNI au niveau transfert, mais par un protocole de niveau service (SIP, RTSP, ...) ou via un portail d'accès (WEB).
2. **Traduction Service – Ressources.** C'est la requête entre le fournisseur de service et l'opérateur comportant les paramètres de QoS techniques. Les données sont formatées sous forme XML pour un transport du type « Web Services » ou sous forme d'AVP pour un transport par le protocole DIAMETER [14].
3. **Pilotage et configuration des équipements.** La QoS est ici explicitement configurée dans les équipements participant à l'acheminement des données. La configuration est effectuée soit équipement par équipement (3a) à l'aide des protocoles² appropriés soit en utilisant une signalisation³ réseau / réseau (3b).

2 COPS, H.248, XML transporté en SSH, Beep (RFC Netconf) voire SNMP ou simplement l'envoi de commandes propriétaires (CLI) en telnet.

3 PNNI pour l'ATM, NSIS ou RSVP pour l'IP

4. **Signalisation centralisée.** La demande de QoS émise par l'utilisateur est portée par une signalisation avec décision centralisée soit par le traitement de service⁴ (4a) soit par le traitement de ressource⁵ (4b). L'équipement peut éventuellement utiliser une signalisation différente pour communiquer avec la plate-forme de gestion centralisée. Par exemple, dans l'architecture de COPS [15], le routeur reçoit une demande via RSVP puis communique avec le PDP en COPS.
5. **Signalisation distribuée.** La demande de QoS émise par l'utilisateur est portée par une signalisation distribuée sur tout ou partie des équipements présents sur le chemin suivi par les données. Les niveaux traitement de service et traitement de ressource ne sont pas impliqués dans le traitement de la demande. Tout se passe au niveau de l'équipement. Le rendu final de la QoS est assuré par la signalisation qui synchronise tous les équipements.

A partir de ces 5 éléments de communication, les différentes architectures de commande de ressources sont classées en 4 familles :

- **Commande Centralisée** utilisant les interactions 1, 2 et 3. Elle regroupe les modèles basés sur une fonction d'admission d'appel qui admet ou rejette une nouvelle session, ou basés sur la gestion de politique de règle d'ingénierie à appliquer dans le réseau sous forme de configuration. La même entité de traitement de ressources contrôle l'ensemble des réseaux dont elle a la charge,
- **Commande Distribuée** utilisant les interactions 1, 2 et 3. C'est une variante de la première famille où la décision est morcelée entre plusieurs partitions de réseaux qui sont sous le contrôle d'entités de traitements de ressources différentes. Ces dernières coopèrent afin de délivrer une réponse cohérente et de bout en bout,
- **Signalisation Centralisée** utilisant les interactions 4 (a et b) et 3a. Elle s'apparente à la famille par commande (centralisée ou distribuée) qui est ici sollicitée via une signalisation usager / réseau pour le traitement de service. Le traitement de ressource est lui sollicité après interception de la signalisation usager /réseau,
- **Signalisation Distribuée** utilisant uniquement l'interaction 5, c'est la signalisation classique, qui suit le chemin emprunté par les données, appelée souvent « On-Path Signalling ».

Lorsque les interactions entre entités fonctionnelles ne suivent pas le chemin emprunté par les données, elles sont qualifiées de méthode « Off-Path » par opposition au « On-Path » de la signalisation distribuée. Les deux familles de commandes sont aussi connues sous le vocable de méthode « PUSH » car elles envoient les configurations dans les équipements, alors que la signalisation centralisée est qualifiée de méthode « PULL » car ce sont les équipements eux-mêmes qui viennent chercher leur configuration.

4.4. Surveillance et Mesure

Dernier processus impliqué dans la gestion de la qualité de service, l'après-vente repose essentiellement sur la mesure, qu'elle soit opérée sur un segment ou de bout en bout. Ces mesures sont exploitées de deux façons différentes : pour la surveillance et la vérification des engagements de la garantie de la qualité de service et comme technique à part entière pour garantir la qualité de service.

4 SIP, RTSP ...

5 RSVP, NSIS ...

D'un côté le client souhaite disposer de moyens de vérifier les engagements de QoS de l'opérateur. De l'autre, l'opérateur a besoin d'outils pour surveiller et mesurer le trafic sur ses réseaux afin d'ajuster la configuration de ses équipements et prouver sa bonne foi en cas de litige. Les mesures sont effectuées soit depuis le terminal du client soit à partir de sondes installées aux points stratégiques du réseau. Les équipements du réseau peuvent également participer à ces mesures. La mesure et la collecte de ces informations sont standardisées pour l'ATM via les flux d'OAM « Operating And Maintenance » [16] et pour l'IP via les messages « NetFlow » [17]. Le système d'information de l'opérateur exploite ensuite ces mesures pour ajuster, au besoin, les configurations des équipements fautifs dans la livraison du service.

Les mesures permettent aussi de déterminer si le réseau est en situation de congestion ou non. L'admission d'appel peut alors s'appuyer sur la mesure de l'état des liens ou de la connectivité entre deux points d'entrée/sortie du réseau, pour déterminer si un nouveau flux peut être accepté.

5. L'admission d'appel

L'admission d'appel ou « Call Admission Control (CAC) » tient une part prépondérante dans la gestion de la QoS. En effet, une fois l'approvisionnement du réseau effectué, et si l'opérateur n'a pas opté pour une gestion de son réseau par surdimensionnement, il est nécessaire de contrôler le nombre de sessions traversant le réseau. Si un conduit a été dimensionné à 10 Mbit/s, la fonction d'admission d'appel doit garantir que ce conduit ne sera jamais utilisé au-delà de ces 10 Mbit/s et cela quel que soit le nombre ou la nature des sessions qui empruntent ce conduit. Cependant, même dans les architectures basées sur une gestion par classe de service, il est nécessaire de disposer d'un contrôle d'admission lors de l'approvisionnement des ressources du réseau. Cette fonction est cependant « noyée » dans l'ingénierie de trafic.

5.1. Définition

Les différents blocs fonctionnels (voir figure 7) manipulent des données de référence (DR) correspondant à des états relativement stables, cependant mis à jour régulièrement, et des données temporaires qui sont des contextes liés à une demande du client, à la commande/livraison (DC) ou à l'utilisation (DU), ou liés à sa consommation. Les contrôles d'admission (au sens le plus général) consistent à comparer les données de la demande client (DC ou DU), aux données de référence (DR) :

- service, débit, QoS, adresses demandées, comparés à l'état du contrat client (profils), à l'état du compte...,
- adresse, débit, QoS demandée, comparés à l'état prévisionnel (EP) ou opérationnel (EO) des ressources, à la souscription et à l'utilisation.

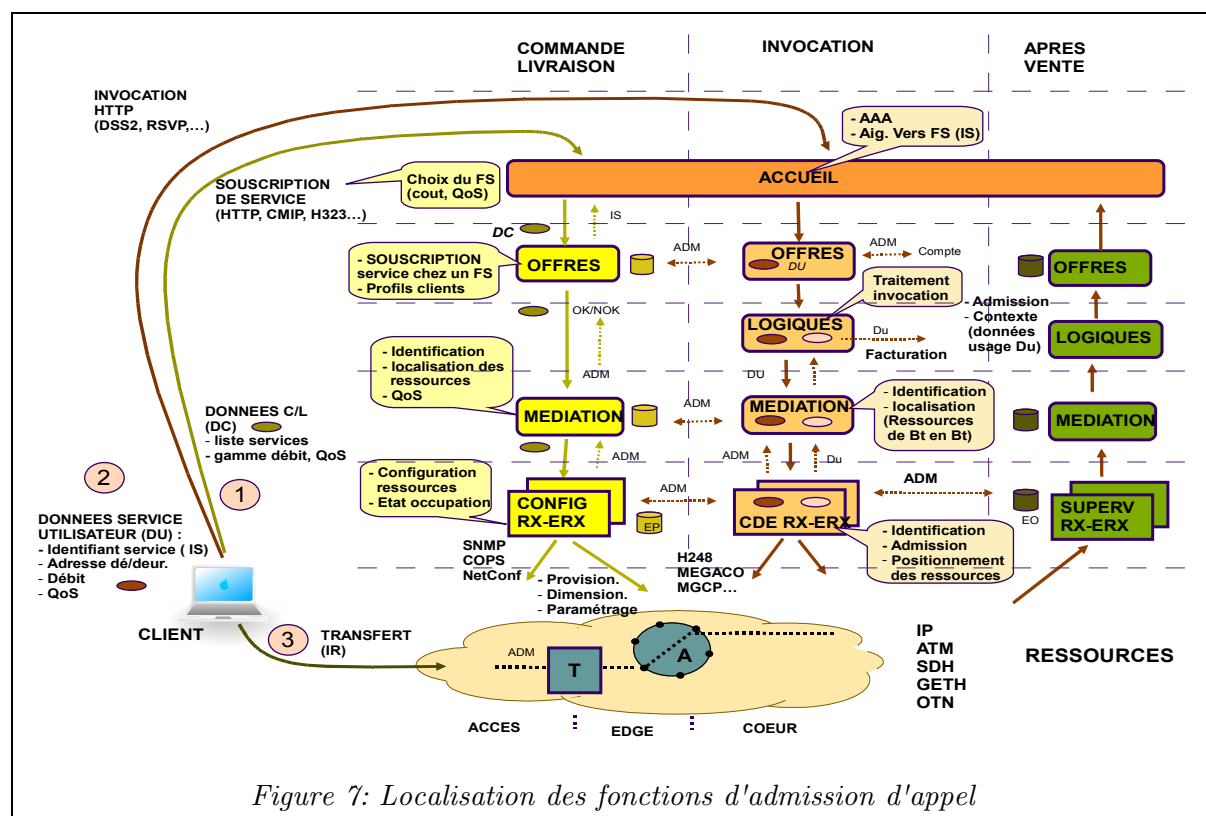


Figure 7: Localisation des fonctions d'admission d'appel

L'état **prévisionnel** des ressources concerne la capacité **utilisée** par rapport à la capacité **maximum** de la topologie du réseau :

- la capacité utilisée au vu simplement des déclarations du client : état **déclaré**,
- ou la capacité utilisée compte tenu d'engagements du client et du fournisseur : état **assuré**.

L'état **opérationnel** des ressources est issu de mesures, d'observations : ce sont donc des états **estimés**. La comparaison à des états déclarés ou estimés conduit à une **QoS relative**, alors que la comparaison à un état assuré permet une **QoS absolue**. Il apparaît donc clairement que la nature du contrôle d'admission conditionne directement la QoS et sa garantie.

Le contrôle d'admission, au sens large, intervient dans tous les niveaux et à différentes échelles de temps : lors des **souscriptions** et de l'approvisionnement du réseau et par rapport à l'état d'occupation des ressources (EP, EO) et lors de l'invocation. Il se décompose :

- au niveau **service** : contrôle que le débit et la QoS requis sont compatibles avec le profil du client, le nombre de sessions autorisées et à l'état de son compte,
- au niveau **médiation d'accès aux ressources** : vérifie que les adresses, le débit et la QoS sont compatibles avec l'état et les capacités des ressources de bout en bout (terminal, accès, sous réseaux, cœur, ...) y compris à la frontière entre opérateurs,
- au niveau **commande des ressources** : contrôle que le débit et la QoS requis sont compatibles avec l'état de référence des ressources, déclaré, ou estimé, ou assuré dans le réseau de l'opérateur,
- au niveau **ressources** : gestion par priorité et police des flux, le cas échéant. Cette police est implicite d'une QoS absolue et donc d'un mode « **assuré** ». Mais elle peut également être appliquée à des flux « Best Effort », pour donner un minimum de qualité aux flux courants.

5.2. CAC de niveau service

Dans le processus de commande / livraison, une nouvelle souscription peut être rejetée essentiellement pour deux motifs : commercial (compte client non approvisionné, facture impayée, dépôt de garantie, ...) ou technique. Dans ce dernier cas, c'est la non éligibilité de l'accès du client au service qui implique le refus. En effet, si le lien de raccordement de l'installation du client au réseau de l'opérateur ne peut supporter le service dans de bonnes conditions i.e. avec une QoS donnée, il est préférable de refuser la souscription à l'offre de service. Sinon, l'opérateur devra supporter au niveau service après-vente un client insatisfait alors même que techniquement, l'accès au service n'est pas possible.

Une fois souscrit, le contrôle d'admission est alors effectué aux différents niveaux du traitement de service au sein du processus d'invocation. Tout d'abord, après authentification, l'utilisateur se voit proposer uniquement les offres de services auxquelles il a accès via son abonnement. Ensuite le contrôle d'admission va vérifier que la demande de l'utilisateur correspond bien à ce qu'il a souscrit : en nombre de sessions simultanées (e.g. 2 communications téléphoniques) et en qualité demandée (e.g. accès aux chaînes TV en haute définition).

Il est important de ne pas confondre la CAC entre les deux processus car l'échelle de temps n'est pas la même. Si l'approvisionnement peut se contenter d'un temps de réaction relativement lent (plusieurs minutes si nécessaire), l'admission d'appel dans le processus

d'invocation doit s'effectuer très rapidement, de l'ordre de quelques millisecondes, afin d'être le plus transparent possible pour l'utilisateur.

5.3. CAC de niveau réseau

Durant le processus de commande / livraison, l'approvisionnement du réseau est amené à vérifier si les ressources physiques des équipements et des liens de transmission sont disponibles à l'aide de fonction d'admission d'appel. Mais, c'est certainement dans le processus d'invocation que la fonction d'admission d'appel prend tout son sens et où elle est la plus connue. Elle vise essentiellement à déterminer si il est possible d'admettre une nouvelle session dans le réseau par rapport à un niveau de QoS exigé. L'algorithme d'admission doit contrôler que les liens logiques approvisionnés dans le processus de commande / livraison ne sont pas utilisés au delà de leur capacité afin de les protéger et garantir ainsi la QoS.

5.4. CAC de niveau transfert

Une fois la session admise, il reste à vérifier la conformité du flux par rapport au contrat négocié entre l'utilisateur et l'opérateur via le fournisseur de service. Seul le processus d'invocation est concerné par ce contrôle d'admission qui est effectué au sein même des équipements du plan de transfert. Au plus fin, il opère au niveau cellule pour l'ATM ou au niveau paquet pour l'IP. Cette admission d'appel est assurée par la fonction élémentaire de police présente dans les équipements. Les paramètres, débit et gigue, utilisés par la fonction de police pour déterminer si un paquet ou une cellule est conforme, doivent être choisis avec la plus grande précaution possible. Trop stricts, ils protègent le réseau contre les rafales mais risquent de détruire à tort des cellules ou des paquets admissibles. Trop souple, ils sont plus tolérants aux variabilités des sources, mais risquent de laisser des rafales ou des paquets non conformes pouvant mettre à mal les files d'attente de l'équipement ou du réseau. Dans le cas de l'ATM, deux techniques applicables à la couche d'adaptation AAL5⁶ permettent de détruire soit la fin d'une trame AAL5, « Partial Packet Discard (PPD) », soit toute la trame AAL5, « Early Packet Discard (EPD) », lorsqu'il n'y a plus de place dans les files d'attente pour stocker le paquet.

6 ATM Adaptation Layer 5 a été standardisée pour transporter des flux de données et est utilisée pour transporter les paquets IP

6. Les architectures standardisées

Dans ce chapitre, les différents modèles d'architecture pour la gestion de la QoS sont passés en revue en mettant en exergue leurs avantages et défauts respectifs.

6.1. Le modèle ATM

6.1.1. Les capacités de transferts

L'ATM [18] se voulant un mode de transfert multi débit et multi service, a placé la gestion de la qualité de service au cœur de son architecture. Plusieurs mécanismes ont été introduits afin de rendre une QoS la plus fiable possible. En premier lieu, la séparation en conduit (VP) et circuit (VC) virtuel permet de séparer les trafics et ainsi les protéger mutuellement les uns des autres. La mise en cellule de longueur fixe et courte permet en outre un multiplexage fin des trafics permettant de garantir un délai et une gigue très faibles pour les services temps réels et conversationnels. L'ATM étant une technologie multi service, la QoS est garantie à travers les capacités de transfert ATM [19] ou « ATM Transfert Capabilities (ATC) ». Elles sont aux nombres de 3 à 7 selon que l'on considère les ATC de base avec ou sans leur subdivision et selon les standards de l'ITU-T ou ceux de l'ATM Forum [20] :

- **DBR (CBR)** pour « Deterministic (Constant) Bit Rate » concerne les applications à débit constant. Le contrat de trafic associé comporte deux paramètres essentiels : le débit (Peak Cell Rate - PCR) et la gigue (Cell Delay Variation - CDV). C'est le plus haut niveau de garantie en terme de QoS mais également le plus contraignant tant pour le client qui doit scrupuleusement respecter son contrat de trafic que pour l'opérateur qui ne peut compter sur un multiplexage statistique pour optimiser les ressources de ses réseaux. Elle est utilisée par les applications nécessitant une QoS stricte sur le délai et la gigue, c'est-à-dire celle de la classe de service Voix.
- **SBR** pour « Statistical Bit Rate » concerne les applications à débit variable. Elle a été déclinée en VBR-rt et VBR-nrt pour « Variable Bit Rate » « real-time » et « non real-time » à l'ATM Forum. Plus compliquée à utiliser, il faut déclarer à la fois un débit maximal (PCR), un débit soutenu (SBR), une gigue et une taille maximale de rafale (nombre de cellules émises consécutivement). Prévue pour les applications utilisant des codecs à débit variable, son ingénierie de trafic fut très longue à élaborer et finalement utilisée que par les opérateurs tant les paramètres à fournir sont complexes à caractériser pour le client. Elle est utilisée pour les applications à débit variable de la classe de service Vidéo,
- **ABT** pour « ATM Block Transfert » permet de réserver des ressources grâce à des cellules spécifiques dites de « Resource Management -RM » avant l'émission des données. Deux modes sont possibles selon que les données attendent le retour de la réservation « ABT/DT – Delayed Transmission » ou sont émises immédiatement « ABT/IT – Immediate Transmission ». Cette ATC n'est présente que dans la recommandation de l'ITU-T. Elle permet de transmettre des données s'accommodant d'un débit variable tout en garantissant durant le transfert du bloc de cellules un niveau de QoS équivalent au DBR. Elle est utilisée pour le transfert de données en mode garanti telle que les applications de la classe de service Donnée,
- **ABR** pour « Available Bit Rate » est la variante de l'ABT à l'ATM Forum. Elle exploite les mêmes cellules RM mais pour obtenir du réseau le débit disponible ou un nombre de crédits plutôt que d'effectuer une réservation absolue,

- **UBR** pour « Unspecified Bit Rate » uniquement spécifiée à l'ATM Forum ne garantit rien en matière de QoS. C'est le pendant de la classe de service « Best-Effort » pour l'ATM. Cette ATC est généralement associée au transfert de données dans un réseau surdimensionné ou conjointement avec la destruction sélective des cellules tel que le « Partial Packet Discard – PPD » ou l'« Early Packet Discard – EPD » afin d'éviter les congestions dans les commutateurs ATM.

Les ATC permettent d'affecter une classe de trafic différente selon la nature même de l'application utilisant ces connexions. Les capacités de transfert DBR, ABT/DT et ABT/IT offrent une garantie de QoS absolue alors que les ATC VBR, ABR et UBR n'offrent qu'une garantie de QoS relative.

En complément de ces capacités de transferts, l'ITU-T a standardisé un type de cellule dédié à la maintenance « Operating And Maintenance - OAM » [16] pour surveiller et mesurer la QoS des connexions. Deux types d'OAM sont disponibles selon que les cellules opèrent de bout en bout, OAM-F5, ou sur un segment du réseau, OAM-F4. Elles permettent en effectuant des mesures au sein même de la connexion (VC ou VP) de vérifier que la QoS est bien respectée au niveau de l'utilisateur à l'aide des cellules OAM-F5, ou sur un tronçon du réseau à l'aide des cellules OAM-F4.

À ces mécanismes de bases du plan de transfert, s'ajoutent les algorithmes d'admission d'appel pour l'approvisionnement des VP, pour le brassage ATM. Chaque équipement, ou la plate-forme de gestion, exécute une admission d'appel en fonction de l'ATC demandé, du débit et des ressources disponibles sur les brasseurs ATM lors de la création du VP. L'approvisionnement du réseau est ainsi contrôlé pour éviter toute surcharge entraînant une dégradation de la QoS.

6.1.2. La signalisation

En complément de la signalisation usager / réseau « User Network Interface - **UNI** » [21] permettant la création à la demande d'un circuit virtuel, VC, l'ATM Forum a standardisé la signalisation **PNNI** « Private Network-Network Interface » [22] pour gérer l'établissement des circuits virtuels dans et entre réseaux. Elle permet d'étendre à l'ensemble des réseaux ATM la fonction de commutation de circuit virtuel « Switch Virtual Circuit - SVC ». La signalisation PNNI intervient en deux temps. Premièrement, au niveau approvisionnement durant le processus de commande / livraison, les commutateurs PNNI échangent des informations de topologies au sein de leur communauté appelée « Peer Group » pour établir un routage à QoS [23]. Ces informations contiennent les adresses ATM joignables ainsi que les liens avec leurs capacités que la signalisation pourra utiliser ultérieurement. Appelés « Trunk Group », ils sont définis par l'opérateur lors de la configuration des équipements et permettent de partitionner et segmenter le réseau en fonction des besoins de l'opérateur.

Le protocole de signalisation de l'ATM PNNI a été conçu dès le départ avec la notion de hiérarchisation. Au sein d'un « Peer Group », un commutateur, en général désigné et configuré comme tel par l'administrateur réseau, est élu comme « Leader ». Il est alors chargé d'agréger les données de topologie du « Peer-Group » auquel il appartient comme étant un simple commutateur ATM. Chaque commutateur acquiert la vision précise de la topologie de son « Peer Group » de niveau n et la vision agrégée des autres « Peer Group » de niveau $n+1$. Une deuxième instance de PNNI est alors exécutée avec les autres « Leader » au sein du « Peer Group » de niveau $n+1$. Tant que la demande de connexion reste à l'intérieur du même « Peer-Group », il n'y a pas de changement de hiérarchie. Si une demande de connexion traverse plusieurs « Peer-Group » différents, son traitement sera

exécuté à différents niveaux de la hiérarchie. Dans le réseau où se situe la source, la demande est traitée au niveau 1, puis le commutateur membre du « Peer-Group » de niveau supérieur va propager la demande en changeant de niveau hiérarchique. Seuls les commutateurs du « Peer-Group » de niveau 2 seront impliqués dans le calcul de la route et dans l'établissement du chemin. L'établissement de la connexion changera de niveau hiérarchique jusqu'à atteindre un « Peer-Group » hiérarchique auquel appartient la destination. Ensuite, chaque commutateur de niveau hiérarchique supérieur à 1 va effectuer une demande dans son « Peer-Group » inférieur afin de calculer plus précisément la connexion.

Ensuite, durant le processus d'invocation, un utilisateur demande la création d'un SVC via la signalisation UNI. Le premier commutateur détermine le meilleur chemin possible pour atteindre la destination en fonction des critères de QoS (ATC + paramètres correspondants) demandés et de la connaissance de la topologie du réseau. Chaque commutateur n'a plus qu'à effectuer l'admission d'appel pour vérifier que les ressources sont bien disponibles. Cette technique, basée sur le « Source Routing » ou routage à la source, est plus efficace que de laisser chaque commutateur décider du chemin à prendre, car le calcul de la route optimale n'est effectué qu'une seule fois. Appartenant à la famille des signalisations distribuées, PNNI est utilisable non seulement durant le processus d'invocation mais il peut également servir au sein du processus de commande / livraison à l'approvisionnement du réseau. Une partie de la configuration de base de PNNI est aussi partie intégrante de ce processus.

Protocole de routage dynamique et adaptatif, il est simple à mettre en œuvre, réagit dynamiquement aux modifications du réseau et évite les boucles de routage. Hiérarchique et multi niveau, il permet l'agrégation d'adresses et donc un meilleur dimensionnement des tables de routage. Autre aspect positif, PNNI s'adapte sans problème à la croissance du réseau et prend en compte les qualités de services dans le calcul du meilleur chemin entre deux points distants. PNNI permet donc d'établir des circuits virtuels dont la QoS est garantie d'une part par l'ATC attachée à la connexion et d'autre part par l'admission d'appel et la réservation effectuée par tous les commutateurs participants à la connexion. Malgré ces propriétés de passage à l'échelle grâce à la hiérarchisation et son excellente gestion de la QoS, le réseau ATM et la signalisation PNNI n'ont jamais réussi à s'imposer : d'une part parce qu'ils avaient perdu d'avance la bataille de l'adressage face à l'IP et d'autre part parce que cette technologie est arrivée en concurrence directe avec la technologie MPLS promulguée par l'IETF.

6.2. Les modèles de l'IETF

La gestion de la QoS ne fait pas l'unanimité au sein de l'IETF. Entre le respect de la « Net neutralité » et le respect du modèle en couche, un certain nombre de groupes de travail ont vu le jour au fur et à mesure des études à partir des années 1990. C'est dans l'ordre chronologique qu'ils sont présentés ci-après.

6.2.1. Modèle Intserv

Le groupe de travail « Intserv » a standardisé une architecture de gestion de la QoS dans l'Internet « **Internet Services** » [24] basée sur deux nouvelles classes de services : « **Guaranteed Service - GS** » [25] et « **Controlled Load - CL** » [26]. Le service GS garantit une transmission sans perte avec un délai maximal pour un débit donné si l'application respecte son contrat. Le service CL garantit seulement un transfert de données équivalent à un réseau peu chargé. La négociation de service s'effectue à l'aide du protocole de réservation « **RSVP** (ReSerVation Protocol) » [27]. Il fait partie de la famille des commandes par signalisation distribuée et opère au sein du processus d'invocation. Basé sur un paquet IP particulier il permet de réserver des ressources, essentiellement de la mémoire

dans les files d'attente plus qu'un débit, au long des chemins à la manière des cellules RM de l'ATM. Cependant, et contrairement à l'ATM, le chemin n'est pas pré-établi en raison de la nature non connectée de l'IP, ce qui rend le protocole plus complexe. De ce fait, le protocole RSVP impose un rafraîchissement régulier des réservations en cours sous peine de les voir disparaître afin de vérifier en permanence que la réservation porte bien sur le chemin courant. A priori « stateless », les routeurs doivent cependant conserver en mémoire l'état des différentes réservations et en particulier les temporisations liées aux rafraîchissements. Séduisant pour des petits réseaux, le protocole s'est rapidement avéré non extensible et non applicable à des réseaux de grande taille et en particulier à l'échelle de l'Internet. Le nombre d'états à maintenir ainsi que le nombre de messages de rafraîchissement ne supportant pas le passage à l'échelle. Plusieurs corrections ont été apportées au fil du temps afin d'essayer de corriger ces défauts, comme la possibilité de rafraîchir simultanément un groupe de réservations.

6.2.2. Modèle DiffServ

A l'opposé de Intserv qui préconise un modèle basé sur la gestion de la QoS par session, le groupe de travail DiffServ pour « **Differentiation Services** » a développé un modèle de gestion par classe de trafic [28] où seul les agrégats de flux se voient traités différemment. Simple à mettre en œuvre, il offre une gestion différenciée par classes de services en gérant par priorités les files d'attente des routeurs correspondant à chaque classe de trafic. Il n'est cependant pas question de garantir une QoS absolue comme dans IntServ ou avec certaines ATC de l'ATM, mais juste d'améliorer le modèle « Best-Effort » [29] en proposant une garantie de QoS relative. Il définit deux services : L'« **Expediting Forwarding - EF** » [30] et l'« **Assured Forwarding - AF** » [31]. Le service EF garantit un délai et une gigue minimums mais ne s'engage pas sur un débit ni sur un taux de perte. Il est bien adapté à la classe de service Voix. Le service AF garantit un taux de perte minimal mais ne s'engage pas sur le délai ou la gigue. Il est bien adapté pour les applications de transfert de données. Ce mode a donné lieu à division : AF1 à AF3 et subdivision de AF11 à AF13 jusqu'à AF31 à AF33. De plus, les RFC ne définissent que le comportement d'un routeur, « **Per Hop Behaviour - PHB** » sans définir un comportement de bout en bout ou au moins à l'échelle d'un réseau. Le routeur d'entrée du réseau, « ingress node », identifie les flux à l'aide de un à cinq paramètres : adresse IP destination et source, port destination et source et protocole. Une fois identifiés, les paquets sont ensuite marqués avec un code appelé « DiffServ Code Point - DSCP ». Ce marquage sert ensuite pour les routeurs, y compris le premier, à aiguiller les paquets dans la bonne file d'attente. Le champ DSCP est véhiculé à la place du champ TOS (Type Of Service) dans l'entête IPv4. Il porte théoriquement sur 8 bits autorisant ainsi 256 priorités différentes. Cependant, le codage des classes de trafic est limité à la partie EXP, soit sur 3 bits seulement ; ce qui donne 8 classes possibles. La valeur 0x0 est réservée pour désigner la classe « Best-Effort » par souci de compatibilité. DiffServ ne réserve pas de bande passante, mais définit le comportement de chaque classe en cas de congestions. Ainsi, si les files d'attente sont pleines, les données des classes les moins prioritaires sont détruites en premier afin de laisser de la place pour les paquets plus prioritaires. Si la destruction des données dans les files d'attente inférieures n'est pas suffisante, les paquets de la file en surcharge sont également supprimés. Le comportement est alors identique à une gestion sans DiffServ. Donc, si le réseau laisse entrer plus de flux dans une classe de trafic donnée, il n'y a aucune garantie de QoS possible. Comme l'IETF ne standardise que des protocoles, les RFC n'ont jamais précisé comment mettre en œuvre les deux services rendant leurs implémentations aussi variées qu'aléatoires.

Pour pallier au problème de surcharge dans une classe de trafic, le groupe de travail a également préconisé l'usage d'une fonction de « **Bandwidth Broker** » [32] afin de gérer, à l'échelle d'un réseau, la bande passante allouée dans chaque classe en fonction des sessions. Véritable admission d'appel, il appartient à la famille des commandes de ressources centralisées. Les architectures à base de « Bandwidth Broker » proposées par Qbone [33] ou dans [34] souffrent cependant de deux problèmes. D'une part, bien que proposées, les interfaces permettant aux applications de formuler leur requête de QoS n'ont jamais été standardisées empêchant une adoption par la majorité des services. D'autre part la difficulté d'acquiescer et maintenir à jour la topologie du réseau ne permet pas de fiabiliser les décisions prises par les algorithmes d'admission d'appel. En cas de pannes ou de re-routage, la topologie n'étant pas à jour, l'admission s'opère sur des données qui ne sont plus à jour. DiffServ appartient donc à la famille des commandes centralisées pour la partie « Bandwidth Broker » et se cantonne au processus de commande / livraison pour la gestion des files d'attente par priorités. Il ne connut jamais de succès auprès des opérateurs du fait de sa relative complexité de mise en œuvre.

Finalement, de DiffServ, seule la gestion par classe est restée. Récemment, l'IETF [3] a cependant précisé les conditions d'utilisation de DiffServ et sa mise en œuvre en fonction du nombre et du type de classe de services qu'un opérateur souhaite déployer.

6.2.3. Modèle IntServ associé au modèle DiffServ

En partant du principe que IntServ n'est applicable qu'en périphérie et que DiffServ convient bien pour la gestion des agrégats, l'idée consiste à associer les deux techniques afin de disposer d'une architecture globale de gestion de la QoS. Les routeurs du réseau cœur sont configurés avec des classes de services (file d'attente par priorité, police) pour lesquels une réservation logique de débit est effectuée par ingénierie de trafic. Les routeurs de périphérie recevant une requête RSVP réalisent alors une admission d'appel sur le débit logique associé avec la classe de service sollicitée par la requête de QoS. À nouveau, ce modèle ne fonctionne que si le réseau cœur est surdimensionné ou si une fonction de « Bandwidth Broker » est disponible. Il est également possible de configurer les PIB du PDP avec les classes de services DiffServ pour une utilisation conjointe avec l'architecture de COPS.

6.2.4. COPS

Comme RSVP était jugé non applicable à l'échelle de grand réseau, un nouveau groupe de travail appelé COPS « Common Open Policy Specification » [15] est parti du constat que RSVP fonctionnait bien en périphérie pour essayer d'améliorer le protocole [35]. Appartenant à la famille des commandes par signalisation centralisée, il est basé sur deux entités fonctionnelles : le « Policy Devision Point - **PDP** » et le « Policy Enforcement Point - **PEP** » [36]. Le PEP, qui réside dans un routeur, recevant une demande de QoS sous forme de signalisation RSVP sollicite le PDP afin de savoir quelle politique appliquer pour ce nouveau flux. Le PDP cherche dans sa base de données quelle est la politique de QoS la plus appropriée à appliquer pour cette demande. Une fois la réponse reçue, le routeur applique la politique fournie par le PDP en configurant les ressources appropriées (classification, police, marquage et file d'attente). Le protocole RSVP est alors terminé dans le routeur et la réponse directement renvoyée à la source. La communication entre le PEP et le PDP est assurée par le protocole COPS du nom éponyme du groupe de travail. Les échanges sont en réalité basés sur la synchronisation de base de données. En effet, le groupe de travail a défini une base de données pour contenir les politiques de règles de QoS à appliquer : « Policy Information Base - **PIB** » [37]. Les PIB sont à COPS ce que les MIB sont à SNMP. Le PDP contient l'ensemble des PIB de tous les PEP qu'il gère. Chaque PEP ne conserve que sa vue

locale au travers de sa PIB. En plus de ce mode « PULL », un mode dit « PUSH » permettait de configurer un PEP depuis un PDP sans utiliser le protocole RSVP [38]. COPS appartient à la famille des commandes de ressources centralisées et utilisables dans le processus d'invocation (mode « PUSH » et « PULL ») ou dans le processus de commande / livraison (mode « PUSH » uniquement).

A nouveau, seul le modèle de PIB pour DiffServ [39] a été standardisé, laissant aux opérateurs le soin de décider quelles sont les politiques à utiliser. Or, la garantie de la QoS dépend exclusivement des règles contenues dans les PIB. Ces politiques se sont révélées très complexes à déterminer pour être réellement adoptées. Finalement, le mode de configuration « PUSH » a souvent été placé en concurrence de SNMP. De ce fait, le protocole COPS n'a pas connu les faveurs de l'IETF et a été peu implémenté par les constructeurs d'équipements.

6.2.5. L'ingénierie de Trafic (MPLS-TE)

À la fin des années 2000, l'IETF a lancé le groupe de travail sur la commutation de label « Multi Protocol Label Switch - **MPLS** » [40] fédérant les propositions de plusieurs constructeurs dont Cisco et 3Com. Les motivations d'origine étaient d'accélérer le routage qui devenait problématique à la fois à cause de l'augmentation du nombre de préfixes annoncés par BGP et à la fois par la complexification de l'algorithme de routage (Longest Match Prefix) dû à la suppression des classes d'adresses [41]. Le principe de base repose sur la commutation de label inséré entre la couche IP et le protocole de transport (MPLS Shim Header) [42] en place et lieu du traditionnel routage bond par bond des paquets IP en fonction de l'adresse de destination. Chaque routeur possède une table de commutation de label qui est mise à jour manuellement ou grâce au protocole de distribution des label : « Label Distribution Protocole - **LDP** » [43]. Le principe consiste à marquer un chemin, appelé « Label Switch Path - **LSP** » ou tunnel MPLS, depuis un nœud d'entrée « ingress node » jusqu'au nœud de sortie « egress node » soit en suivant les tables de routage soit en explicitant les routeurs par où le LSP doit passer. Ces routeurs sont appelés « Label Edge Router - **LER** » et sont chargés de labelliser les paquets en entrée et de supprimer le label en sortie du domaine MPLS. Les labels sont ensuite commutés lorsque les paquets transitent par un routeur MPLS, appelé « Label Switch Router - **LSR** ». MPLS supporte également l'empilement des labels, « Label Stacking », autorisant de fait une hiérarchisation des labels. Contrairement à l'ATM qui ne supporte que 2 niveaux de hiérarchisation i.e. le VPi et le VCi, MPLS autorise jusqu'à 256 niveaux de label. La contrepartie est l'augmentation de la taille des paquets et le risque de fragmentation dû à cet accroissement.

Se voulant indépendant du plan de transfert, MPLS a été décliné sur de nombreuses technologies. Ainsi, l'en-tête MPLS est recopiée dans le champ VPI/VCI des cellules ATM afin de rendre la technique compatible avec la commutation ATM [44]. Les paquets IP arrivant sur le nœud d'entrée sont labellisés suite à leur classification selon une table de « Forwarding Equivalent Class - FEC ». Une FEC contient toutes les informations nécessaires à l'identification des paquets depuis un préfixe réseau jusqu'au cinq paramètres habituels (adresses source et destination, ports source et destination et protocole). Un moment évoqué comme pouvant être utilisé durant le processus d'invocation pour établir des LSP session par session à la manière de la commutation ATM, MPLS demeure avant tout une architecture dédiée à l'approvisionnement du réseau, trouvant sa place dans le processus de commande / livraison. Elle appartient à la famille des signalisations distribuées. Par la suite, de nombreuses extensions ont été ajoutées à MPLS :

- La prise en compte de DiffServ [45] qui permet de construire des LSP compatibles avec les classes de service de DiffServ soit en initialisant un LSP par classe (L-LSP)

soit en recopiant les 3 bits EXP du champ DSCP dans l'en-tête MPLS (E-LSP) afin de gérer les paquets dans des files d'attentes séparées.

- La protection des LSP avec la possibilité de créer des LSP de secours. En cas de panne, le routeur qui détecte l'erreur a la possibilité d'utiliser un LSP de secours pour court-circuiter localement ou globalement le routeur défaillant. Cette technique appelée « Fast-ReRoute » permet de re-router le trafic d'un LSP à un autre en moins de 50 ms, comme pour la SDH, garantissant une excellente qualité de service.
- La possibilité d'utiliser MPLS à l'inter-domaine via MPLS-BGP [46].
- La prise en compte des flux de maintenance afin de surveiller les LSP.
- La prise en compte du multicast avec la possibilité de créer des LSP point à multipoints.
- La généralisation de MPLS en GMPLS, qui prend en compte différentes couches physiques dont le transport optique avec la possibilité de commuter des lambdas. GMPLS est ainsi devenu le protocole de configuration et de signalisation des commutateurs optiques.

Très rapidement, la possibilité de gérer la QoS par ingénierie de trafic est apparue comme une recommandation forte dans les évolutions de MPLS. L'idée était non seulement pouvoir spécifier par où passe le LSP, mais en plus de pouvoir réserver de la bande passante sur les routeurs que le LSP emprunte. Une bataille entre le protocole RSVP-TE [47], extension de RSVP à l'ingénierie de trafic et CR-LDP [48], une extension de LDP, s'est engagée pour standardiser le protocole permettant d'établir des LSP contraints. Au final, seul RSVP-TE est resté. Il est associé aux protocoles de routage à ingénierie de trafic comme OSPF-TE [49] et IS-IS-TE [50]. En effet, RSVP-TE étant « state less », c'est le protocole de routage qui indique la bande passante restante sur chaque lien en plus de leur état. Chaque routeur prend alors une décision locale pour déterminer si il dispose de suffisamment de bande passante pour acheminer le LSP ou non. En cas d'erreur, un mécanisme de retour en arrière « Crank-Back » permet d'essayer d'autres chemins dans l'établissement du LSP. Connue sous le vocable de MPLS-TE [51], les opérateurs disposent aujourd'hui de toute une boîte à outil pour construire et gérer leur réseau à l'aide LSP.

Plus récemment, MPLS-TE a été également adapté à DiffServ. Ainsi, MPLS-DiffServ-TE [52] permet de réserver de la bande passante pour des LSP (E-LSP ou L-LSP) affectés aux classes de services de DiffServ. La réservation s'effectue selon deux modes. Le modèle conservateur « Maximum Allocation Model -MAM » [53] qui n'autorise pas un LSP à prendre de la bande passante dans un LSP de priorité inférieure qui en disposerait ; le modèle dit des poupées russes « Russian Doll Model - RDM » [54] qui permet le chevauchement des LSP de priorités supérieures sur les LSP de priorités inférieures. Le routage des paquets dans les LSP est effectué directement à partir du champ DSCP sans nécessiter de configurer une entrée dans la table de FEC.

Enfin, comme le calcul du placement de tunnel est complexe pour un routeur qui n'a pas toujours la visibilité de tout le réseau (cas du multi aire d'OSPF ou du multi région de IS-IS), un nouveau groupe de travail s'est formé qui a pour objectif de standardiser un protocole de requête « Path Computation Protocol - PCEP » [55] de calcul de placement de tunnel entre un « Path Computation Client -PCC » et un « Path Computation Element - PCE » qui fait office de serveur de calcul [56]. Le PCC et le PCE peuvent être localisés indifféremment dans un routeur ou un serveur externe. Limitée pour l'instant à l'intra domaine, des extensions sont en cours pour utiliser l'architecture PCE à l'inter-domaine.

Le groupe travail MPLS, après bientôt 10 ans d'activités est toujours actifs à l'IETF preuve que cette technique a les faveurs de la communauté scientifique, des industriels et des opérateurs. L'architecture MPLS est certainement l'une des plus abouties en matière de gestion de la QoS dans les réseaux IP. La signalisation MPLS permet au final de recréer un mode connecté par commutation de label au dessus de l'IP tout en gardant la souplesse du routage bond par bond. Le seul reproche qui peut lui être fait, est qu'elle n'est pas disponible pour toutes les infrastructures de transport comme les réseaux d'accès radio.

6.2.6. Inter Domaine

Si l'ATM Forum avec le protocole PNNI propose une solution efficace et extensible via le concept de hiérarchisation, l'IETF ne propose rien en matière de gestion de QoS à l'inter-domaine. En effet, le protocole BGP-4 [57] ne permet que d'échanger des vecteurs de distances pour constituer des chemins d'AS. De plus, afin de rester extensible et applicable à l'échelle de l'Internet, BGP agrège les routes qu'il a appris avant d'envoyer ses annonces à ses voisins. Les métriques et les filtres utilisés par BGP reflètent plus les politiques des opérateurs que la QoS offerte et supportée par les réseaux. Les chemins d'AS obtenus décrivent uniquement des routes dites « Best-Effort ». Ainsi, un AS de bordure n'a que peu de choix sur l'ensemble des chemins d'AS possibles pour joindre une destination donnée et aucune possibilité de choisir une QoS.

Plusieurs projets européens ont proposé des alternatives à BGP afin de pallier à cette déficience. AQUILA [58] a développé **BGRP** [59] comme une extension de BGP permettant de réserver des ressources sur les liens inter-domaines. MESCAL [60] a tout d'abord étudié la cohérence de la QoS entre réseaux afin de garantir une continuité des classes de services [61]. Puis, il a développé **qBGP** [62] qui permet de calculer des chemins d'AS différents en tenant compte de métriques de QoS échangées par BGP. Présenté à l'IETF, qBGP a finalement été transmis à l'IRTF car considéré comme trop en avance par rapport au besoin courant de stabilisation et de fiabilisation du protocole BGP. Plus récemment, Cisco [63] a proposé une modification dans l'implémentation de BGP afin de pouvoir exécuter plusieurs instances de BGP entre deux routeurs au sein de la même connexion. Les métriques étant différentes d'une session à l'autre, il devient alors possible d'obtenir un routage différent en fonction de paramètres de QoS.

6.2.7. NSIS

RSVP étant malgré tout quelque peu utilisé, l'IETF a décidé de le rajeunir et surtout pallier aux problèmes inhérents du modèle IntServ. « Next Steps in Signaling - NSIS » [64] pourrait n'être qu'un RSVP v2, mais le modèle en deux couches de signalisations le rend beaucoup plus puissant et flexible. En premier lieu, la couche de signalisation générique « General Internet Signaling Transport - GIST » [65] assure l'acheminement des paquets de signalisation depuis une source jusqu'à une destination via un ensemble de routeurs. Contrairement à RSVP, tous les routeurs ne sont pas obligés de supporter le protocole. Comme RSVP, cette signalisation est « On-Path » dans le sens où elle suit le même chemin que les données même si tous les routeurs n'y participent pas. Au dessus de GIST, plusieurs signalisations dites applicatives « NSIS Signalling Layer Protocol - NSLP » sont en cours de standardisation, dont la gestion de la QoS [66] et des FiveWall. NSIS est partie intégrante de la famille des signalisations distribuées et utilisable dans le processus d'invocation. De plus, le groupe de travail a défini comment utiliser NSIS dans le cadre d'un domaine DiffServ contrôlé par une fonction d'admission d'appel et de préemption du trafic [67]. Cette dernière permet de classer également NSIS dans la famille des signalisations centralisées.

Malgré toute la bonne volonté du groupe de travail, les principaux documents sont toujours à l'état de brouillon après plus de 4 ans de travail. GIST et la signalisation QoS-NSLP sont en cours de révision par les instances administratives de l'IETF et devraient bientôt sortir en RFC. Il restera en dernier recours à vérifier dans la pratique si NSIS est réellement applicable aux réseaux de grande taille ou si il ne souffre pas des mêmes défauts que son ancêtre RSVP, ce qui semble être le cas au vu de sa complexité croissante.

6.2.8. Flow Aware et PCN

La technique dite du « Flow-Aware » [68] considère les flux de données comme étant un tout indivisible qui une fois admis dans le réseau vont être acheminés sans qu'il soit nécessaire de leur appliquer un traitement particulier. Par contre, il est nécessaire de contrôler le nombre de flux entrant dans le réseau afin que celui-ci reste fluide. Elle peut être qualifiée de « Early Flow Discard - EFD » par analogie à la technique de l'« Early Packet Discard (EPD) » de l'ATM ou toutes les cellules d'un paquet AAL5 sont détruites si il n'y a pas de place dans les files d'attente du commutateur, puisque tous les paquets composant le flux sont détruits en entrée du réseau si il n'y a pas assez de ressources pour acheminer le flux en question. Comme pour l'EPD, l'EFD nécessite d'être déployé dans tous les routeurs du réseau puisque les ressources ne sont vérifiées que localement. Une variante consiste à effectuer une mesure entre le nœud d'entrée et de sortie du flux pour vérifier l'état du réseau et ainsi appliquer l'EFD uniquement en entrée du réseau.

Le nouveau groupe de travail de l'IETF « Congestion and Pre-Congestion Notification - PCN » procède de façon similaire à la technique du « Flow Aware ». Son cahier des charges est la protection des flux en cas de congestion avérée ou imminente. La solution est basée sur l'admission des nouveaux flux ou la terminaison des flux sur les routeurs d'entrées d'un domaine DiffServ en se basant sur des indications ou mesures de congestion ou pré congestion [69]. Le groupe de travail en est à ses débuts et a prévu de standardiser, outre l'architecture, les conditions et techniques de collecte des indicateurs de congestions, leur encodage et leur transport ainsi que les techniques d'admission d'appel. Cette architecture utilise à la fois le processus d'après-vente pour les mesures et la collecte des indicateurs, et le processus d'invocation pour l'admission d'appel dans le plan de transfert. Sans pouvoir préjuger du résultat final, il faut veiller à ce que la solution proposée ne souffre pas des mêmes problèmes que la technique de l'« Early Congestion Notification - ECN » alliée au « Generic Flow Control - GFC » de l'ATM [70]. En effet, dans le cas des réseaux à haut débit, les liens de raccordement se comportent comme une mémoire en plus des files d'attente des routeurs. Le temps nécessaire pour détecter une congestion et la notifier vers la source est souvent plus grand que la période d'émission de la source. La notification intervient alors trop tard i.e. quand la source n'a plus rien à émettre.

Les deux techniques souffrent de problèmes communs. Il est en effet très difficile de déterminer à coup sûr la fin de la session. Comme ces deux techniques se veulent indépendantes des applications, il est impossible de connaître le débit nécessaire à l'écoulement du flux, ni sa durée, ce qui rend l'algorithme d'admission particulièrement délicat à mettre au point. Enfin, dans le cas d'applications multi flux, telle que la visioconférence (un flux audio + un flux vidéo), il est également impossible de corréliser les flux entre eux et donc le risque de ne pas admettre un flux mais pas l'autre est grand, rendant ainsi l'application inutilisable.

6.3. Les modèles du NGN

Fin des années 1990 et début des années 2000, est apparu comme une nécessité pour les opérateurs de rationaliser leurs différents réseaux afin de réduire les coûts d'investissement et d'exploitation. Comme l'ATM peinait à s'imposer et que l'IP devenait de plus en plus prépondérant, il fallait trouver une architecture à la fois flexible et commune afin de fédérer l'ensemble des réseaux et des services. Connu sous le vocable de NGN « Next Generation Network », cette architecture est répartie des fondamentaux décrits dans l'outil OURAf en séparant organiquement les 3 plans : service, contrôle et transport. Le problème majeur est que beaucoup d'architectures et de solutions ont été présentées en se revendiquant compatibles avec le NGN sans être compatibles entre-elles. Aujourd'hui, 2 modèles ont émergé de cette période de gestation : l'IMS « Internet Multimedia SubSystem » standardisé à l'ETSI et le NGN standardisé à l'ITU-T [71].

6.3.1. L'IMS

L'ETSI, aux travers du 3Gpp pour le mobile et de TISPAN pour les réseaux fixes, a standardisé l'IMS [72] comme réseau fédérateur pour transporter tout type de contenus multimédia. Le premier service supporté est bien sûr la voix, en mode natif sur IP en utilisant le protocole d'établissement de session, « Session Initiation Protocol - SIP », et en mode émulation pour s'adapter et s'interconnecter avec le réseau téléphonique commuté (RTC). L'ETSI est cours de standardisation de l'IMS IP TV afin de fournir des services vidéos (TV, Vidéo à la demande, magnétoscope dans le réseau, ...). L'architecture est construite autour du cœur IMS (voir figure 8) qui assure les fonctions d'authentification, d'autorisation, de comptage (AAA) et de routage des appels.

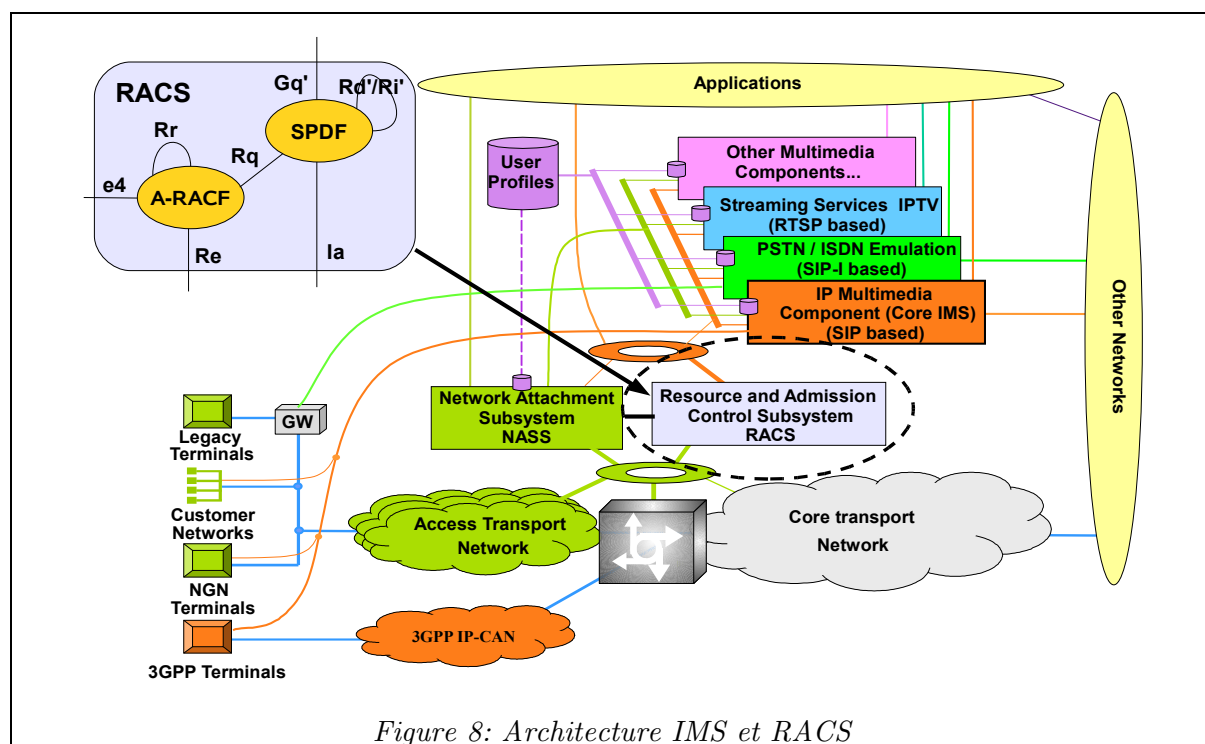


Figure 8: Architecture IMS et RACS

Plusieurs fonctions additionnelles ont été standardisées dont la gestion de QoS. Elle est appelée « Resources Allocation and Control Subsystem - **RACS** » [73] pour les réseaux fixes

et « Policy Decision Function - PDF » pour les réseaux mobiles. Ces entités assurent les fonctions d'admission d'appel et de configuration du nœud d'accès au cœur IMS. De plus, le RACS est en liaison avec le sous-système de gestion de l'attachement au réseau, « Network Attachment Sub-System - NASS » afin de collecter les informations sur la ligne de l'utilisateur : l'adresse IP qui lui a été dynamiquement affectée ainsi que le profil de la ligne et du service souscrit.

La RACS se décompose en deux éléments : le « Service Policy Decision Function - **SPDF** » qui de façon analogue au PDF du mobile vérifie la conformité de la demande de QoS et configure le nœud d'accès et le « Access Resource Allocation and Control Function - **A-RACF** » qui réalise la fonction d'admission d'appel et de réservation de ressources proprement dites. Les interfaces entre toutes ces fonctions ont été standardisées de sorte à pouvoir être intégrées facilement. En particulier, l'interface Gq'⁷, au dessus du SPDF, rend son utilisation possible par d'autres entités que le cœur IMS, comme des serveurs d'applications. Le RACS appartient à la famille des commandes centralisées et est utilisé par le processus d'invocation.

Malgré l'engouement actuel pour l'IMS de la part des opérateurs pour renouveler leurs services conversationnels, la gestion de la QoS n'en demeure pas moins partielle et incomplète. Tout d'abord, elle est trop spécifique au protocole SIP. Les paramètres à fournir au RACS via l'interface Gq' sont extrapolés du protocole SIP. Certaines données, comme la liste des codecs ou l'identifiant de session rend l'interface finalement peu adaptée aux autres services comme les « Web Services ». Ensuite, l'ETSI s'est focalisé uniquement sur les réseaux d'accès (mobile ou fixe) délaissant les réseaux cœurs (considérés comme surdimensionnés) et les relations à l'inter-domaine pourtant fondamentaux pour les communications internationales. Bien qu'une interface entre cœur IMS ait été identifiée, les études sur l'interconnexion ont été repoussées à plus tard. Récemment, la relation entre RACS et réseau domestique, i.e. la passerelle résidentielle, a été identifiée et le travail de standardisation de l'interface est en cours de définition. De même, la version 2 du RACS inclut la gestion des réservations de ressources pour les flux diffusés. Enfin, l'acquisition de la topologie du réseau d'accès est laissée à l'appréciation des constructeurs. Libre à eux de proposer une méthode pour renseigner ces informations. Seul le couplage avec le NASS permet de récupérer dynamiquement l'association entre adresse IP et ligne physique de l'accès.

6.3.2. L'ITU-T

En parallèle de la standardisation de l'IMS, l'ITU-T a démarré ses travaux pour normaliser le NGN [74]. Au sein de cette architecture, l'ITU-T a normalisé une fonction d'admission d'appel et de contrôle de la QoS appelée « Resource Allocation and Control Function - RACF » [75]. Elle se démarque de l'ETSI (voir figure 9) dans le sens où elle matérialise clairement la séparation entre le niveau 2 (dépendant de la technologie de transport) et le niveau 3 (indépendant de la technologie). Chaque niveau est également séparé entre les fonctions qui réalisent la QoS (fonction C, P, M, F et S du modèle OURAf) et la commande de ces mêmes fonctions. Ainsi le « Policy Decision Functional Entity - PD-FE » contrôle le « Policy Enforcement Functional Entity - PE-FE ». Le PD-FE réalise les fonctions de la Médiation de Ressources du modèle OURAf : La gestion de l'inter-domaine, le choix des réseaux, l'admission d'appel dans le domaine et la configuration des équipements de périphérie. De même, le « Transport Resource Control Functional Entity - TRC-FE » qui pilote le « Transport Resource Enforcement Functional Entity - TRE-FE » est en charge du

7 L'interface Gq' est basée sur le protocole DIAMETER et des AVP spécifiques à la demande de QoS. Elle est compatible avec l'interface Gq utilisée par le 3Gpp pour les demandes de QoS.

niveau 2 pour l'admission d'appel et la configuration des équipements. L'inter-domaine a été pris en compte dès le départ et plusieurs scénarios sur la coopération entre RACF sont évoqués en annexe du document Y.2111. Dans la deuxième mouture de la norme en cours de finalisation, le multicast est également supporté. Enfin, l'interface avec les applications a été conçue dans l'optique d'être vraiment multiservice et non contrainte à l'utilisation du protocole SIP.

Bien que non définies en terme de protocoles, plusieurs interfaces ont été identifiées à la fois pour la configuration des équipements et pour l'acquisition des informations de topologies du réseau. Cependant, la norme Y.2111 reste une spécification fonctionnelle ne décrivant que les interfaces avec l'extérieur. A l'image de beaucoup de normes de l'ITU-T, peu de plates-formes sont disponibles.

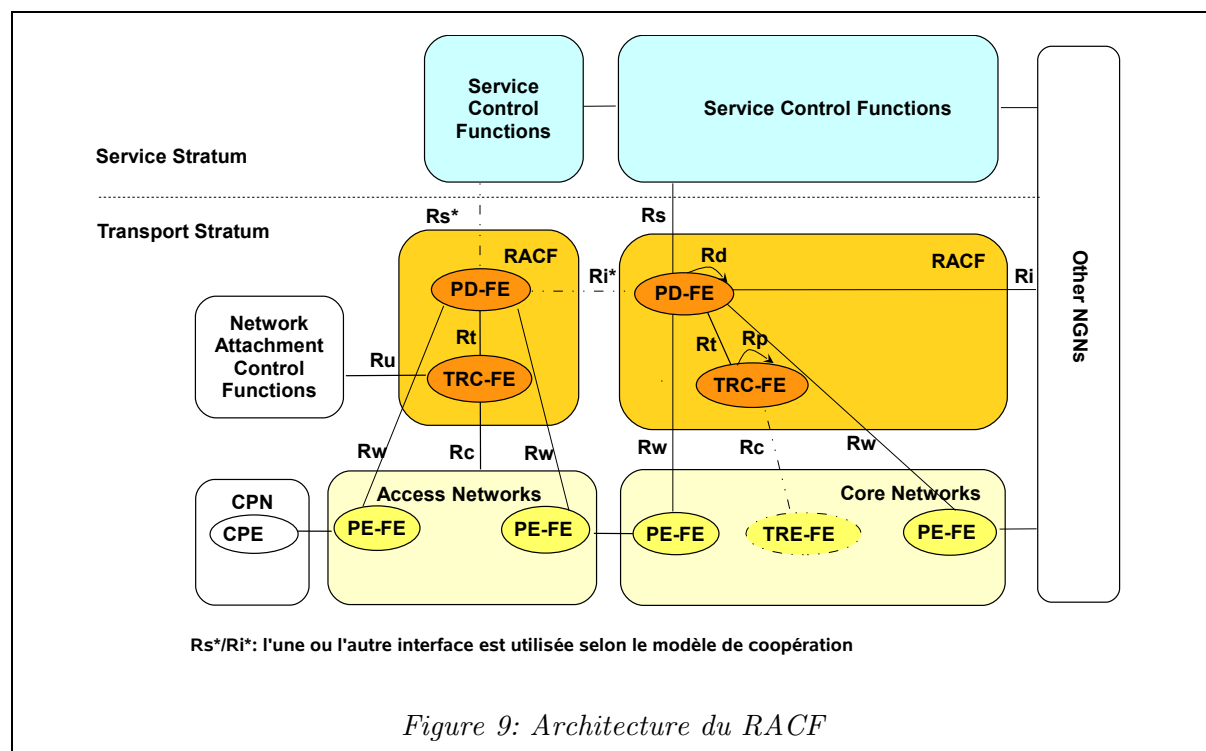


Figure 9: Architecture du RACF

7. Analyses des solutions

7.1. Synthèse de l'état de l'art

L'ensemble des architectures et des mécanismes de gestion de la QoS n'offrent cependant pas de solutions satisfaisantes en particulier si elles doivent s'appliquer aux différentes technologies sous-jacentes et garantir une QoS de bout en bout. Seule l'ATM répond en partie à ces critères, mais comme c'est une technologie en perte de vitesse, elle ne représente pas la solution d'avenir utilisable par les opérateurs qui sont passés majoritairement à l'IP. La figure ci-dessous cartographie l'ensemble des architectures évoquées précédemment.

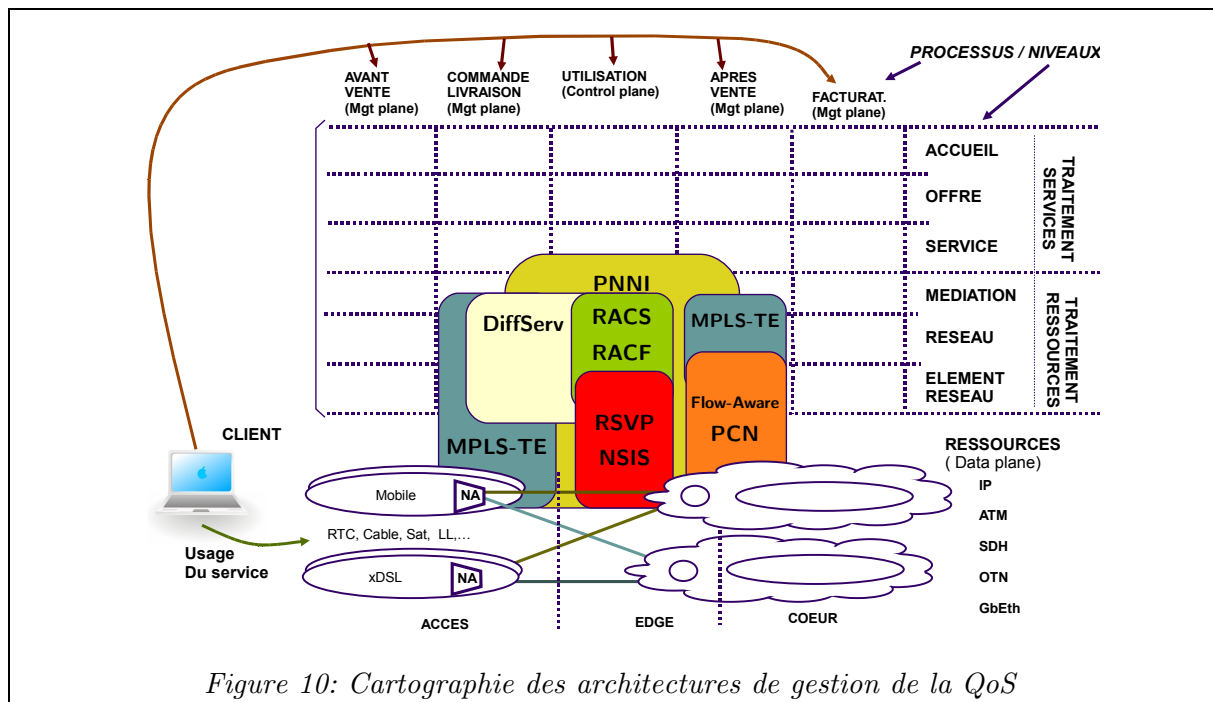


Figure 10: Cartographie des architectures de gestion de la QoS

Les solutions abordées sont également difficilement transposables aux réseaux de grande taille et supportent difficilement le passage à l'échelle, RSVP en étant peut-être la figure emblématique. L'inter-domaine et les relations entre opérateurs ne sont pas ou très peu prises en compte. Un opérateur est ainsi dans l'impossibilité de pouvoir offrir une continuité de la garantie de qualité de service à ses clients dès lors que les connexions doivent traverser des réseaux qu'il ne contrôle pas directement.

Côté classe de service, chaque organisme de standardisation a défini ses propres classes de services sans chercher à être compatible ni réutiliser ce qui a déjà été produit. Ajouté le manque d'API pour les programmeurs d'applications leur permettant de les exploiter, les CoS restent finalement le seul apanage des opérateurs lorsqu'ils approvisionnent leur réseau. Seul le récent RFC 5160 [76] issu des travaux de Levis [61] aborde à titre informatif la continuité de classe de service pour fournir une QoS de bout en bout.

Le surdimensionnement reste également une technique très en vogue, mais qui risque très rapidement de n'être plus supportable par les opérateurs avec l'arrivée des raccordement en fibre optique offrant des débits montants allant jusqu'à 100 Mbit/s voire plus en GPON.

7.2. Limitations des solutions actuelles

Les solutions proposées ne permettent pas d'élaborer une architecture complète pour gérer la QoS de bout en bout, mais offrent n architectures partielles ou mécanismes qu'il faut assembler tel un « légo ».

7.2.1. Acquisition et synchronisation de données

Les architectures basées sur la commande centralisée implémentant des fonctions d'admission d'appel et de réservation de ressources sont qualifiées de solution « Off-Path ». En effet, la demande de QoS ne suit pas le même chemin que les données. Afin de pouvoir réaliser une gestion efficace de la QoS, il est nécessaire de connaître la topologie du réseau contrôlé ainsi que le chemin que va emprunter les données. Les différentes implémentations utilisent en général une base de données qui contient les informations de topologie ainsi que les réservations en cours sur les différents liens et équipements. Il faut donc pouvoir alimenter cette base de données automatiquement surtout dans le cas du contrôle de réseaux de grande taille. De même, en cas de modification de ce chemin (re-routage dans les réseaux IP, pannes, ajout, ...), la fonction d'admission d'appel doit impérativement être informée de ces nouvelles données afin de mettre à jour sa base de données et exécuter de nouveau les algorithmes d'admission d'appel et de réservation de ressources sur cette nouvelle topologie. Hormis l'écoute des protocoles de routage des réseaux IP, il n'existe pas de solution satisfaisante pour acquérir automatiquement et maintenir à jour les topologies des réseaux. Cela rend, d'un point de vue opérationnel, l'utilisation de ces architectures complètement caduque.

7.2.2. Paramétrages et lien avec les applications

Si un opérateur souhaite garantir une QoS relative, une gestion par classe de service et priorité des files d'attente est suffisante. Par contre, il est vain de croire qu'une solution de gestion de la QoS peut être découplée des applications dès lors qu'il souhaite garantir un niveau de qualité de service absolue. Pour cela, il est indispensable de pouvoir synchroniser la demande de QoS avec le début et la fin de l'application. En effet, la solution doit pouvoir gérer les erreurs engendrées par des pannes, des arrêts brutaux, ... sous peine de devoir gérer des réservations fantômes dans le réseau. Détecter les débuts et fins d'activités au niveau TCP est très délicat et encore moins faisable au niveau UDP. Il faut donc remonter dans les couches protocolaires, pour arriver finalement au niveau de l'application. En outre, il est impératif de disposer de paramètres fiables et correspondant aux données de service e.g. débit et durée d'une vidéo, nombre de flux, ... La difficulté vient à la fois pour les applications de spécifier correctement les paramètres de QoS associés à leur besoin réel et pour le réseau d'identifier les flux d'une application donnée. À nouveau, la classification à base du quintuplet (adresse ip source et destination, port source et destination, protocole) ne fonctionne pas dans tous les cas, notamment lorsque les paquets IP sont fragmentés ou lorsque les applications sont situées derrière une fonction de NAT. Le recours à des marqueurs simples, sans toutefois devoir gérer des agrégats comme avec le champ DSCP, est indispensable. Même si IPv6 n'a pas de gestion native de la QoS comme IPv4, il permet aux applications de marquer leur flux à l'aide du champ « flow label » transporté dans l'entête des paquets IPv6. Cet identifiant va simplifier la classification des paquets, puisqu'il permettra d'identifier plus sûrement les flux d'une application. Finalement, il manque aux applications, et plus particulièrement aux développeurs, une API de programmation permettant de façon standardisée d'émettre une requête de QoS vers le réseau.

7.2.3. Performance et passage à l'échelle

Toute solution se doit d'être extensible et pouvoir être appliquée à des réseaux de grande taille. Dans l'inventaire réalisé, seules les techniques garantissant une qualité de service relative peuvent prétendre à une application à toute échelle de réseau, y compris à l'échelle de l'Internet. Les signalisations telles que RSVP et très certainement NSIS ont échoué dans cet exercice. Seul PNNI et MPLS-TE qui ont été conçues dès le départ avec un souci de hiérarchisation, supportent l'extensibilité.

Si l'augmentation de la taille des réseaux est déjà un problème en soi, il faut également compter avec l'augmentation du nombre d'utilisateurs. Ainsi, le nombre de requêtes de QoS par unité de temps va suivre une loi en $\Theta(n^2)$ où n désigne le nombre d'utilisateurs. Même si toutes les applications n'ont pas besoin de QoS particulière, et donc peuvent se contenter du mode « Best-Effort », le nombre d'états à maintenir dans le réseau va également croître selon la même loi. À nouveau, seules les solutions hiérarchiques abstrayant les niveaux inférieurs comme PNNI et MPLS-TE, permettent de s'affranchir de ce problème en gérant, à un niveau donné, un nombre borné de réservations.

Toujours au chapitre des performances, la configuration dynamique des équipements est et restera certainement un goulot d'étranglement pour les solutions de commandes de ressources centralisées. Réalisées la plus part du temps à l'aide de commandes opérateurs envoyées par telnet, les performances sont très insuffisantes (de l'ordre de 2 configurations par seconde) pour pouvoir être utilisables durant le processus d'invocation. Même si COPS procure théoriquement de meilleures performances, il est peu supporté par les constructeurs. Reste le récent groupe de travail « NetConf » de l'IETF qui si il est bien implémenté⁸ par les industriels pourrait devenir la solution de choix à la configuration dynamique.

7.2.4. Gestion de l'inter-domaine

Dernier problème partiellement résolu à ce jour, l'inter-domaine est l'un des points clés pour le succès du déploiement de la QoS. Si les opérateurs sont dans l'impossibilité d'offrir une continuité de QoS de bout en bout, leurs services à valeur ajoutée resteront cantonnés dans leur réseau sans pouvoir rivaliser avec les ténors de l'Internet que sont Google, MSN, Dailymotion, ... Aujourd'hui seul les accords de gré à gré entre opérateurs pour l'interconnexion de leur réseau sont utilisés pour la configuration de BGP. Cependant, de gros volumes de données transitent sur les réseaux des fournisseurs d'accès sans qu'il soit possible pour ces derniers de facturer la bande passante consommée par ces services, car le modèle de coût vers les usagers est du type forfait fixe et les fournisseurs de contenus ne sont pas directement raccordés au réseau du fournisseur d'accès. De ce fait, les opérateurs doivent supporter des coûts de raccordement importants sans pour autant pouvoir les impacter sur l'abonnement de leurs clients. Une gestion de la QoS à l'inter-domaine permettrait de gérer ces trafics tout en offrant des possibilités de rémunération supplémentaire.

7.3. QoS mythe ou réalité ?

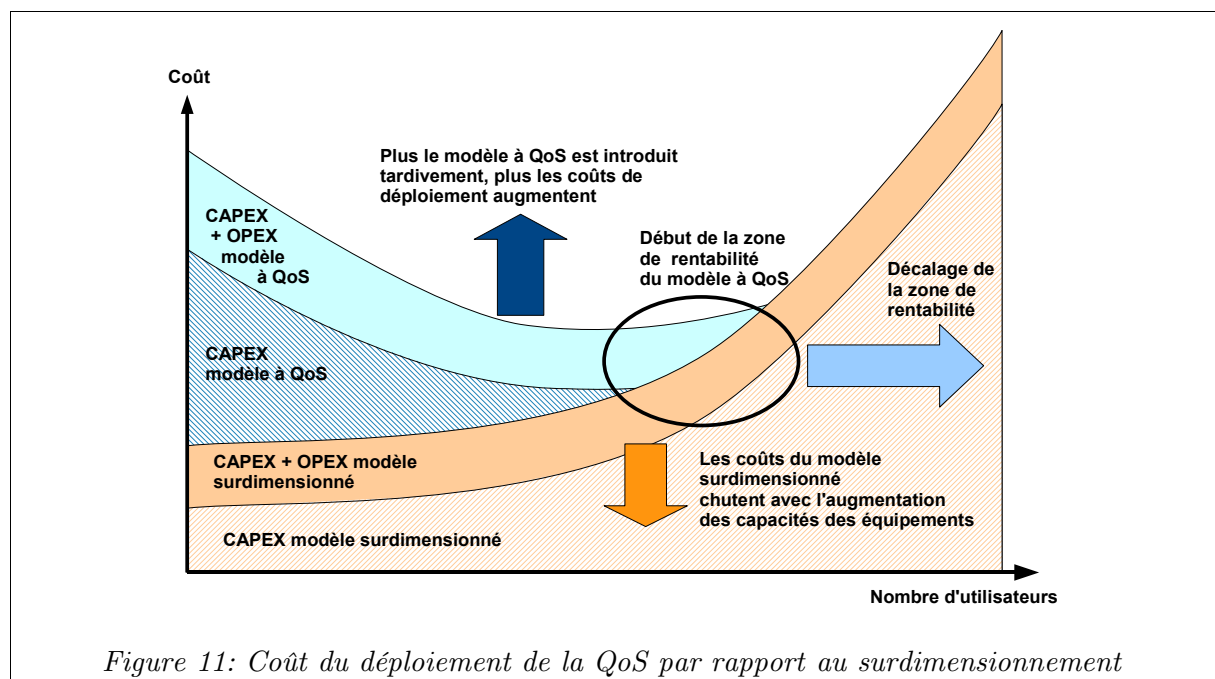
Il ne s'agit pas ici d'effectuer le procès de tel ou tel mécanisme ni de rentrer dans telle ou telle polémique scientifique mais plutôt de déterminer les raisons techniques et économiques qui ont conduit les opérateurs à peu ou pas déployer de mécanismes de gestion de la QoS dans leur réseau. Ce bilan reste aujourd'hui d'actualité et risque de le rester encore pour quelque temps.

8 Dont les performances ne sont pas bridées pour des raisons de sécurité par exemple

7.3.1. Contraintes économiques

Avant même de débattre sur le coût relatif de telle ou telle solution de QoS, un opérateur de réseau se pose d'abord la question du coût d'introduction d'un mécanisme de gestion de la QoS par rapport à augmenter les capacités de son réseau. En effet, les opérateurs opposent d'abord le modèle dit du surdimensionnement (over-provisioning) aux mécanismes de gestion de la QoS de façon générale.

La figure ci-dessous illustre la différence de coût entre le surcoût du surdimensionnement du réseau et le coût d'introduction d'un mécanisme de contrôle de la QoS. La courbe ne fournit qu'une enveloppe (tendance) en fonction du nombre d'utilisateurs et non un chiffrage précis. En effet, ce dernier est trop lié i) aux coûts de fonctionnement propre de chaque opérateur, ii) au type de mécanismes de contrôle de la QoS et iii) aux prix des équipementiers de télécommunication. De plus, il est très dur, voire impossible, d'obtenir les coûts réels auprès des opérateurs. La comparaison tient compte non seulement des coûts d'investissements (CAPEX), mais également des coûts liés à l'exploitation (OPEX).



Dans le cas du modèle surdimensionné, l'augmentation des coûts est majoritairement portée par les investissements (CAPEX) du fait de la multiplication des équipements et des liaisons. Ce facteur explose lorsque l'expansion du réseau nécessite d'augmenter de manière drastique le nombre de liaisons entre équipements induisant des travaux de génie civil. Par contre, la part due aux coûts d'exploitation (OPEX) reste quasiment constante quel que soit le nombre d'utilisateurs. En effet, la technologie restant la même (ex. Routeur IP et connexion GigaBit Ethernet), les OPEX n'augmentent que d'un facteur correspondant à l'accroissement de la taille du réseau.

Les mécanismes de contrôle de QoS ont une structure de coût différente dont l'évolution passe par un optimum. Soit R_{QoS} le rapport entre l'optimisation des ressources du réseau et les capacités totales d'acheminement. Un réseau surdimensionné aura un R_{QoS} inférieur à 20%. L'optimum entre le gain apporté par l'optimisation des ressources du réseau

et le coût inhérent des mécanismes de gestion de la QoS correspond à un R_{QoS} compris entre 60 et 80%. En dessous de ce seuil, la part des investissements n'est pas rentabilisée par le nombre d'utilisateurs. Cela correspond aux coûts fixes d'achat d'une telle solution (qui peuvent correspondre au coût des licences par exemple). Au-delà, les coûts d'investissements dans une solution plus complexe qui augmente nécessairement les coûts d'exploitation ne sont plus rentabilisés par les gains d'optimisation d'usage des capacités du réseau. À noter que le coût total d'une solution avec gestion de QoS doit prendre en compte le coût du réseau non surdimensionné.

En première approximation, le contrôle de la QoS devient intéressant lorsque le seuil de rentabilité de la QoS est atteint. Cependant, plusieurs facteurs viennent interférer avec cette conclusion. En premier lieu, la technologie évoluant et les coûts de fabrication diminuant, le surcoût du surdimensionnement du réseau tend à diminuer avec le temps. L'opérateur américain SPRINT a évalué que le surdimensionnement de son réseau correspondait à anticiper ses investissements de seulement 6 mois. En parallèle, les mécanismes de contrôle de la QoS semblent nécessiter de plus en plus d'investissements et tendent donc à augmenter avec le temps. À cela peut s'ajouter le coût d'introduction de tels mécanismes qui peuvent devenir prohibitifs au fur et à mesure qu'un opérateur développe ses réseaux. Plus le nombre de services déployés augmente, plus les coûts d'adaptation de ces services au mécanisme de contrôle de la QoS augmentent.

La loi de « moore », qui régit l'évolution des semi-conducteurs, peut également s'appliquer aux télécommunications, mais avec des sauts temporels de 5 ans au lieu de 18 mois. Ainsi, le débit d'accès offert aux clients est multiplié par 10 tous les 5 ans. La zone de basculement entre le modèle surdimensionné et le modèle à QoS se voit sans cesse repoussée, différant ainsi la prise de décision qui devient de plus en plus difficile à prendre avec le temps. Cependant, le prochain saut technologique risque d'avoir raison du modèle surdimensionné. Le surdimensionnement des réseaux de collecte et de cœur dépend du débit montant des liens d'accès offerts aux clients, 800 kbit/s maximum en xDSL. La migration progressive vers la fibre optique va multiplier le débit montant par un facteur compris entre 10 et 100 (selon les offres de services) et risque de prendre de vitesse l'évolution des réseaux de collecte et de cœur. Cette migration va impliquer un saut de la zone de croisement entre les deux modèles dû à un décalage vers le haut des coûts du modèle surdimensionné.

7.3.2. Contraintes techniques

La complexité de mise en œuvre de solutions de QoS, ou tout du moins son passage du laboratoire à l'exploitation est également un frein technique à son déploiement. Si certaines solutions fonctionnent très bien sur des réseaux expérimentaux, leur mise en service sur des réseaux de grande taille et en environnement perturbé (Denis de Service, interférence, panne, ...) peut s'avérer laborieux et long à stabiliser. Ceci est une règle assez générale qui ne vaut pas que pour la QoS (ex. la mise au point de la stabilité du protocole BGP à l'échelle de l'Internet). Cependant, un système de contrôle de QoS mal maîtrisé ou mal géré peut s'avérer pire en cas de dysfonctionnement que l'absence du même mécanisme. Ainsi, une plate-forme d'admission d'appel centralisée est immanquablement un point faible supplémentaire dans un réseau, dont la panne peut engendrer l'arrêt complet des services qu'elle est censée protéger.

Tout d'abord, la solution retenue doit supporter le passage à l'échelle i.e. elle doit pouvoir s'appliquer à des réseaux de grande taille et supporter leur évolution. Il a déjà été prouvé que certaines technologies étaient inappropriées comme, par exemple, RSVP. Cependant, le critère de passage à l'échelle a souvent servi d'argument ultime pour envoyer au panier des technologies pourtant fort prometteuses. Même si certaines technologies sont moins extensibles que d'autres, c'est un faux problème que d'opposer systématiquement le

passage à l'échelle pour contrer le déploiement d'une solution de contrôle de la QoS.

Ensuite, certains mécanismes de contrôle de la QoS, en particulier ceux du type centralisé, nécessitent d'être alimentés en données par le système d'information de l'opérateur. La topologie du réseau, la localisation des clients, la description des équipements contrôlés et leur accès (login/passwd) sont autant de données nécessaires à l'exécution d'une fonction admission d'appel. Le souci réside dans la non standardisation des systèmes d'informations qui sont propres à chaque opérateur et à l'absence d'interfaces normalisées pour accéder aux données. Ceci rend l'intégration du mécanisme de contrôle de QoS complexe induisant des coûts supplémentaires, car il est réalisé spécifiquement pour un opérateur et un réseau donné. À cela s'ajoute également l'adaptation du mécanisme aux équipements du plan de transfert du réseau de l'opérateur. À nouveau, même si certains équipements sont très répandus, il s'agit de développements spécifiques.

7.4. Conclusions

Même si toutes les architectures présentées souffrent de défauts, les opérateurs de télécommunications ont cependant déployé des mécanismes de gestion de la QoS.

Tout d'abord, **l'ingénierie de trafic**, qui consiste à établir les règles de déploiement et d'utilisation, est un préalable à la conception d'un nouveau service. Qu'elles soient statiques ou dynamiques, les règles d'ingénierie de trafic peuvent largement subvenir aux besoins de QoS des applications. MPLS commence à être largement déployé par les opérateurs qui voient ainsi la possibilité de gérer efficacement leurs réseaux tout en offrant de la QoS.

L'**admission d'appel** est également utilisée à différents stades. En premier lieu, l'éligibilité du client au service en fonction de la configuration de son point d'accès devient un critère d'admission au service permettant d'assurer un premier niveau de QoS. Ainsi, un utilisateur qui n'est pas suffisamment proche d'un DSLAM se verra refuser toute demande de souscriptions à des offres multi services nécessitant un débit plus élevé que ce que peut supporter sa ligne. Dans un deuxième temps, les opérateurs profitent que certaines plateformes de services embarquent des fonctions d'admission d'appel pour les activer. Les coûts sont ainsi fortement réduits puisqu'il n'est pas nécessaire d'intégrer un nouvel équipement dans le réseau. Cependant, cette gestion mono service empêche toute optimisation et mutualisation de la gestion de la QoS dans le cas du multiservice.

Finalement, c'est au niveau du **plan de transfert** que les opérateurs ont le plus porté leur effort. Que ce soit en utilisant les capacités de transfert de l'ATM (ATC), le marquage des flux avec DiffServ ou le cloisonnement par VLAN, la technique reste la même. Il s'agit de gérer les flux de façon différenciée en leur assignant une priorité différente. Cette gestion des applications par classes de services différenciées se décline dans toutes les partitions du réseau. Sa mise en œuvre reste simple car elle est basée sur la configuration statique des équipements réalisée à la construction du réseau et réajustée peu fréquemment. Les flux sont, soit marqués par la plate-forme de service ou les terminaux, soit après classification par les nœuds d'accès au réseau, à partir de règles d'identifications. Enfin, d'un point de vue des partitions de réseau, les efforts se portent essentiellement sur les points de congestions qu'il n'est pas possible de surdimensionner : la paire de cuivre et l'accès au réseau de collecte pour l'xDSL, l'accès au lien radio pour la téléphonie mobile ou le Wifi, ...

Au final, face aux techniques de surdimensionnement et de multiplexage statistique majoritairement utilisées dans les réseaux d'opérateurs, mais n'offrant pas de QoS ou une toute relative, il ne reste que les techniques d'approvisionnement. C'est le modèle de gestion par classe de service qui est privilégié, soit en utilisant le modèle DiffServ pour une gestion

par priorité simple, soit en utilisant MPLS-TE pour pré réserver et garantir des ressources. Il est donc souhaitable de pouvoir disposer d'une architecture et de recommandations permettant d'assembler les différents mécanismes de gestion de la QoS afin de disposer d'une solution globale, extensible et fiable.

Les manquements identifiables à ce jour et qui vont être développés dans le prochain chapitre, concernent donc :

- Une interface et une API de programmation unique permettant aux applications d'exprimer clairement et simplement leur besoin en matière de QoS,
- Une interface et un cadre unique permettant aux opérateurs de s'échanger des requêtes de qualité de services afin de garantir de la QoS de bout en bout,
- Une définition des objectifs de QoS par technologie pour garantir une continuité de la qualité de service,
- Une synchronisation de la topologie du plan de transfert avec la fonction d'admission d'appel.

III

Contributions aux travaux d'Architectures pour la QoS

« L'architecte, c'est formuler les problèmes avec clarté »

Charles-Edouard Le Corbusier

Cette partie se propose d'exposer les contributions de l'auteur aux architectures de gestion de la QoS en partant de l'analyse effectuée dans l'état de l'art. La démarche scientifique a toujours suivi les mêmes règles au fil de l'évolution de mes travaux de recherches : Comment fournir de la qualité de service à l'aide de mécanismes d'admission d'appel et de commande de ressources dans les réseaux de transfert en mode paquet ?

Quatre axiomes de base ont servi de guide tout au long de ces années de recherches :

- Identifier et séparer le traitement des paquets pour permettre un contrôle des flux de données connexion par connexion,*
- Mettre en forme les flux de données au plus prêt de la source afin d'en faciliter le multiplexage dans le réseau,*
- Réserver des ressources dans les équipements pour chaque application nécessitant une garantie absolue de QoS,*
- Protéger le réseau et les flux de données, en cours d'acheminement, contre les surcharges en effectuant une admission d'appel préalablement à l'introduction dans le réseau de tout nouveau flux requérant de la QoS,*
- Décomposer les solutions spatialement (en fonction des partitions de réseaux concernées) et temporellement (selon les processus de commande / livraison, d'invocation et d'après-vente).*

Les différentes contributions sont exposées à la fois dans l'ordre chronologique de leur étude et à la fois en remontant les couches et les niveaux où elles s'appliquent.

1. Contrôle et Multiplexage par espacement

La technique de l'espacement fût mise au point dans les laboratoires de France Telecom dans les années 90 [77] et [78]. Au début, elle a été conçue comme un moyen de contrôle et de remise en forme du trafic afin de faciliter le multiplexage des flux dans un réseau ATM. Il s'est avéré, par la suite, que l'espacement est une discipline de gestion des files d'attente et une technique de multiplexage à part entière. Les contributions de l'auteur dans le domaine de l'espacement concernent :

- L'étude et la mise en œuvre du multiplexage des cellules à l'aide d'un mécanisme de tri original des heures d'émissions,
- L'étude et la formalisation des règles nécessaires à la régulation sans perte de flux TCP transportés par une connexion ATM.

1.1. Espacement et échéancier virtuel

Les études sur le multiplexage statistique des cellules débutèrent avec la mise au point de la technique ATM et se renforcèrent par la suite [79]. Il était nécessaire, pour la réalisation des premiers commutateurs, de disposer de modèles mathématiques sur le comportement du multiplexage des flux ATM afin de pouvoir dimensionner leurs files d'attente. A titre d'exemple, l'ATC VBR ou SBR repose sur le seul contrôle d'admission pour garantir de la QoS, qui est lui-même issu directement de la modélisation mathématique de la superposition du trafic VBR. Parmi ceux-ci, le modèle composite de file d'attente $M+D/D/1$, qui permet de représenter le multiplexage d'une source périodique avec un flux de poisson, donne comme résultat un flux de poisson [80]. Ce résultat constant a permis de modéliser le comportement du réseau entier lorsqu'il est confronté à des connexions à débit constant. En d'autres termes, le multiplexage de source à débit constant donne un flux de poisson, parfaitement modélisé et prévisible, applicable à l'échelle d'un réseau. Malheureusement, le monde des télécommunications, hormis le réseau téléphonique classique, est loin d'être constitué de sources à débit constant et les autres tentatives de modélisations mathématiques sur des sources plus génériques ont rarement abouti à un résultat pouvant se généraliser facilement à l'échelle d'un réseau. L'espacement des cellules a été inventé en partant du constat qu'il était toujours possible de rendre une source quelconque à débit constant. Le principe repose sur deux fondements :

1. Toute source suit un échéancier virtuel d'émission,
2. L'espacement selon un échéancier virtuel doit être protégé par une fonction de contrôle.

1.1.1. L'algorithme VSA(T, τ)

L'échéancier d'une connexion correspond aux heures théoriques d'arrivées des cellules de l'affluent ATM distantes de la période T_{PCR} déclarée. Chaque cellule arrivant à l'accès est numérotée d'indice k . La cellule arrive à l'heure a_k . L'algorithme d'espacement virtuel [78] calcule son heure théorique d'arrivée « Theoretical Arrival Time » TAT_k à l'aide de la période T_{PCR} comme suit :

$$TAT_{k+1} = \max(TAT_k, a_k) + T_{PCR} \quad \text{avec } TAT_0 = a_0$$

La gigue affectant un affluent ATM, va retarder une cellule - effet de dispersion - et rapprocher la suivante - effet d'agrégation. Soit y_k la gigue entre deux cellules d'une même connexion comme étant la différence de temps entre l'heure d'arrivée réelle de la cellule a_k et son heure théorique d'arrivée TAT_k . Si la cellule est en retard, y_k est négatif, si elle est en avance, y_k est positif. L'espacement virtuel est alors contrôlé en fixant une valeur maximale à

y_k appelée tolérance de gigue et notée τ . Si la source ne respecte pas son contrat de trafic en envoyant trop de cellules dans un intervalle donné, on obtient $y_k > \tau$. Dans ce cas, la cellule est détruite et l'échéancier virtuel TAT_k ne progresse pas : $TAT_{k+1} = TAT_k$. Si la source voit ses cellules retardées ou si elle n'émet rien pendant un intervalle donné, on obtient $y_k < 0$. Dans ce cas, l'échéancier est réinitialisé avec l'heure réelle d'arrivée, $TAT_k = a_k$, pour éviter qu'une source ne mette à profit une période de silence pour accumuler des « crédit » de temps au niveau de l'algorithme.

Cet algorithme rebaptisé « Generic Cell Rate Algorithm - GCRA » a été normalisé avec sa variante le « Leaky Bucket » [81] à état continu dans la recommandation I.371 de l'ITU-T [19] et sert d'algorithme de référence pour mesurer la performance et la conformité des flux ATM dans la recommandation I.356 de l'ITU-T [2]. Toute source peut ainsi se conformer au $VSA(T, \tau)$. Il suffit de calculer la période de référence T de la connexion et la gigue maximale τ de sorte à ce que le flux rentre dans le gabarit de la fonction $VSA(T, \tau)$. Un opérateur voulant protéger l'usage de son réseau exigera une gigue τ faible, ce qui peut imposer une période T élevée et donc un débit élevé pour les applications à débit variable. Au contraire, un opérateur plus tolérant acceptera des applications avec une gigue τ très large ce qui permet d'ajuster la période T , et donc le débit, au plus près du débit « moyen » de l'application.

1.1.2. Du VSA à l'espace

Une fois les heures théoriques d'arrivées obtenues à l'aide de l'algorithme $VSA(T, \tau)$, l'espace réel s'obtient en émettant la cellule selon son TAT plutôt que sur la base de son heure réelle d'arrivée [82]. Ce dispositif est connu sous le nom de contrôleur espaceur dans la littérature [83]. La connexion n'est pas ralentie et ne se voit pas offrir un débit inférieur à celui négocié durant la caractérisation de son trafic (calcul des valeurs de T et τ). Au contraire, tout se passe comme si la connexion était transportée par une liaison louée qui lui était exclusivement réservée avec non seulement garantie du débit négocié, mais également une gigue et un délai faible. Le $VSA(T, \tau)$ associé à l'espaceur réel, permet d'offrir une QoS correspondant à la classe Voix quel que soit le type de connexion.

Cependant, sa mise en œuvre est plus délicate surtout lors du passage d'un échéancier virtuel à un échéancier réel des heures d'émissions des cellules. Comme plusieurs connexions ATM peuvent entrer en conflit pour la même heure de départ, il est nécessaire de pouvoir gérer les conflits d'accès au médium de sortie pour une heure donnée. La problématique réside dans la possibilité de trier des heures afin de connaître la plus petite valeur et ceci dans un temps cellule, soit 2,726 μ s pour un multiplex STM1 à 155,52 Mbit/s.

Deux types de systèmes peuvent être utilisés : commande à l'entrée et commande à la sortie. La première catégorie regroupe les systèmes qui cherchent la place où stocker la cellule dans la mémoire de sortie de l'espaceur [84] et [85]. La mémoire est lue séquentiellement. Une cellule est émise si la case est occupée. Lorsque le pointeur de lecture arrive à la fin, il est rebouclé au début de la mémoire. La valeur maximale de gigue admissible est alors égale à la taille de la mémoire. Les systèmes à commande par la sortie stockent les cellules avec leurs heures théoriques d'émission dans une mémoire commune au fur et à mesure de leur arrivée. Les heures sont alors triées de sorte à ce que la cellule ayant l'heure d'émission la plus petite (la plus proche de l'heure courante) soit émise en premier lorsque son heure d'émission arrive à échéance i.e. inférieure ou égale à l'heure courante. La valeur maximale de la gigue est égale à la valeur la plus petite entre la précision du codage des heures et la taille de la mémoire de stockage. Ils font appel à des comparateurs organisés en systole [86], en registre à décalage [87] ou en arbre binaire [88]. Des composants de type mémoire associative peuvent également être utilisés. La première contribution de ce mémoire est la mise au point

d'un système de tri [89] des heures d'émissions à deux étages qui permet à la fois de minimiser le nombre de portes logiques du composant tout en offrant un système rapide et extensible.

1.1.3. Vers une discipline de service universelle

L'espacement réel est donc une discipline de service des files d'attentes basée sur les heures théoriques d'arrivées des cellules. Elle est universelle car elle respecte les caractéristiques temporelles intrinsèques de l'application. Le débit offert correspond à celui qui a été négocié lors de l'établissement du contrat de trafic, et le délai et la gigue sont minimisés puisque égaux aux valeurs résiduelles d'un étage de commutation en présence de files d'attentes courtes. Certes, il ne permet pas de bénéficier de tout le débit disponible lorsque le réseau est vide ou faiblement chargé comme le permet la classe de service « Best-Effort » ou une simple file d'attente FIFO. Cependant, en terme de QoS, il offre la possibilité de pouvoir implémenter une seule discipline de file d'attente accouplée à une mémoire de grande taille permettant d'accepter des tolérances de giges τ très élevées. En effet, une cellule sera toujours servie en fonction de son heure théorique d'émission, même si elle est arrivée après les cellules des autres connexions et quel que soit le nombre de cellules en attente dans l'espaceur.

1.2. Le multiplexage par espacement

Le VSA s'est avéré être une discipline de service à part entière en même temps que l'idée de l'appliquer au multiplexage de plusieurs entrées d'un même commutateur ou multiplexeur. En effet, le tri des heures d'émissions est non seulement un mécanisme de sélection de la prochaine cellule à émettre sur l'interface de sortie, mais il est de plus indépendant de l'origine des cellules. Ainsi, le tri peut s'effectuer sur des cellules en provenance de différents ports d'entrées. Deux architectures matérielles sont possibles pour implémenter le multiplexage par espacement :

- Soit l'ensemble des cellules en provenance de tous les ports d'entrées peuvent atteindre le port de sortie sans perte (cas des matrices de commutation sans blocage), alors le contrôle espacement peut s'effectuer en sortie,
- Soit le VSA est effectué sur chaque port d'entrée, puis les cellules sont multiplexées en fonction de leur heure d'émission.

Si le premier cas de figure peut sembler plus simple à mettre en œuvre, car il suffit de disposer d'un seul VSA en sortie, sa réalisation est toutefois beaucoup plus délicate car il doit pouvoir fonctionner à la même vitesse que la matrice de commutation i.e. $n \times 155,52$ Mbit/s pour n entrées soit un temps d'exécution de $2,726 \mu s / n$. C'est finalement la technologie qui limite fortement le nombre d'entrées supportées (entre 8 et 16).

La deuxième contribution de ce mémoire est un principe novateur de multiplexage selon les heures d'émissions des cellules [90] qui fonctionne au débit de l'interface d'entrée la plus rapide et non à la somme des débits d'entrées. Le tri des heures, permettant d'effectuer le multiplexage des cellules, s'effectue en deux temps. Tout d'abord, chaque interface d'entrée effectue le contrôle des affluents ATM selon l'algorithme $VSA(T, \tau)$ associé à chaque connexion puis trie les heures d'émissions théoriques. À chaque temps cellule, chaque entrée présente la cellule estampillée avec la plus petite heure d'émission. Puis, dans un deuxième temps, toutes ces heures d'émissions vont de nouveau être triées afin de déterminer l'heure la plus petite qui désignera ainsi la prochaine cellule à émettre. La gigue de multiplexage de l'ensemble correspond à celle de 2 contrôleurs espaceurs, soit deux étages de multiplexage FIFO à 64 places. La collecte des TAT et des cellules est effectuée à l'aide d'un bus de

multiplexage circulaire actif original. Chaque emplacement physique sur le bus comporte un composant actif qui réalise à la fois la collecte, la circulation des cellules et la comparaison des heures. Après une révolution complète, l'emplacement de l'interface de sortie sur le bus présente non seulement la cellule à émettre, mais également le numéro de l'interface présentant la cellule dont le TAT est le plus petit [89]. Le tri est effectué de proche en proche sur le bus selon la formule $TAT(t) = \min (TAT(t-1), TAT_i)$ où $TAT(t)$ et $TAT(t-1)$ représente respectivement l'heure la plus petite au temps t et $t-1$ et TAT_i représente l'heure de la cellule de l'interface i . Le principe de multiplexage peut se résumer comme étant le calcul du $\min(\min(TAT_k))$. Le bus de multiplexage actif doit respecter deux contraintes :

- Il doit assurer la distribution de l'horloge de référence sur toutes les entrées afin que la comparaison des heures ait un sens. Il faut en effet que toutes les entrées soient synchrones du point de vue du calcul des heures sous peine de comparer des heures dont le référentiel temporel n'est pas le même et obtenir un résultat erroné,
- La fréquence du bus correspond au temps cellule sur le multiplex de sortie divisé par le nombre d'entrées afin que chaque entrée puisse proposer une heure d'émission et émettre une cellule pour l'interface élue. Chaque étage du bus actif étant composé essentiellement d'un comparateur qui commande un sélecteur et d'un registre à décalage, la réalisation matérielle supporte sans difficulté un nombre conséquent d'entrées (de 32 à 128).

Un deuxième principe [91] plus simple permet de s'affranchir de la première contrainte. Dans ce cas, les cellules sont collectées selon un principe FIFO. Un jeton circule sur le bus en place et lieu de la comparaison des heures. La première entrée sur le bus qui a une cellule à émettre prend le jeton. Au cycle suivant, l'entrée émet sa cellule. Si la mise en œuvre est plus simple car il n'y a pas besoin de synchroniser toutes les interfaces d'entrée sur la même horloge, la gigue résultante peut s'avérer plus importante. Les premières interfaces sur le bus sont privilégiées par leur emplacement et peuvent ainsi s'approprier, plus que nécessaire, le médium de sortie, par rapport à la comparaison des heures d'émissions mentionnées précédemment.

1.3. Régulation de TCP par espacement

La couche d'adaptation AAL3/4 était censée fournir une couche de transport des données sur ATM avec garantie de QoS. Devant sa complexité et le semi échec de sa standardisation, l'ITU-T mis au point la couche d'adaptation AAL5, pour lui succéder, qui peut se résumer à une simple mise en cellule des paquets TCP/IP. Seul le dernier paquet porte un indicateur de fin de trame dans l'en-tête de la cellule et un CRC32 sur l'ensemble du paquet transporté dans la payload. Très simple, cette adaptation ne permet pas la garantie de QoS. C'est la simple transposition du mode « Best-Effort » de l'IP à l'ATM. Aussi, à la fin des années 90, les études ont essentiellement porté sur le transport de TCP/IP sur l'ATM tout en lui garantissant de la QoS. Avec les mécanismes de « Back-Pressure » à base de « Generic Flow Control », de l'usage du bit de priorité des cellules « Cell Loss Priority - CLP », de « Partial & Early Packet Discard – PPD, EPD » [92], toutes les tentatives ont plus ou moins échoué. En parallèle, et sur fond de conflit entre l'ATM et l'IP, l'IETF a apporté des améliorations au protocole TCP [93] afin qu'il puisse s'accommoder d'une couche de transport non fiable. Ce qui est le cas de la couche IP. Les mécanismes de « Slow Start », de « Congestion Avoidance », de « Fast Recovery » et de « Fast Retransmit » [94] visent à modifier la gestion standard de la taille de la fenêtre de congestion de TCP « TCP Congestion Window ». Le slow start permet à TCP de faire croître la taille de la fenêtre progressivement (selon une loi exponentielle) jusqu'à rencontrer une perte. Une fois ce point atteint, la taille de fenêtre est ramenée à 0 et « le slow start » repart jusqu'à la

moitié de la valeur atteinte précédemment avant la perte. Ensuite, c'est une simple progression linéaire qui est utilisée par le mécanisme d'évitement de congestion jusqu'à rencontrer à nouveau une perte. Si le plan de transfert n'offre aucune perte, les deux mécanismes s'arrêtent lorsque la taille maximale de la fenêtre est atteinte. En cas de perte, TCP se régule autour d'une valeur moyenne, celle offerte par la connexion, mais en mode tout ou rien et en perdant systématiquement des données. Malgré le consensus établi dans la communauté scientifique, il ne s'agit pas véritablement de régulation au sens système régulé et stabilisé (comme en électronique) mais plutôt d'oscillation de débit avec perte entre 0 et n , aboutissant à un débit effectif de $n/2$.

Dans ce contexte, le contrôleur espaceur démontra très rapidement les effets bénéfiques de l'espacement sur les flux AAL5. Il permettait de « lisser » les flux TCP, dont la caractéristique première était d'être très « bursty ». Les drivers des cartes ATM étaient conçus comme pour les interfaces Ethernet. Dès que le média est disponible, la carte envoie toute la trame au débit du multiplex de raccordement i.e. 155,52 Mbit/s pour un STM-1. Un paquet de 1500 octets produit ainsi une rafale de 32 cellules. Les premiers commutateurs ATM n'implémentant que des FIFO de petite taille, 64 cellules maximum, les connexions paquets TCP produisaient des pertes dans les commutateurs et les connexions étaient impossibles à établir. Le « slow start » et le « congestion avoidance » ne s'appliquant qu'à la gestion de l'émission des paquets, il n'y avait aucun moyen de régler le débit d'émission des cellules. Au fur et à mesure des développements du contrôleur espaceur et de son application au flux TCP [95], un constat s'est rapidement imposé : non seulement l'espacement des cellules régulait le trafic en entrée des commutateurs, protégeant ainsi leur files d'attente, mais il permettait également de réguler le protocole TCP [96].

Lorsque la connexion ne présente pas de perte de cellule, l'émetteur envoie ses paquets TCP/IP jusqu'à ce que la fenêtre de congestion de TCP « TCP congestion window » arrive à saturation (voir figure 12). TCP reste dans la phase « slow start ». L'émetteur doit alors attendre le retour des accusés de réception pour voir de nouveau sa fenêtre d'émission augmenter et donc pouvoir émettre de nouveaux paquets. Si les paquets sont espacés au débit de la connexion, l'accusé de réception des paquets va être reçu par l'émetteur une fois ces derniers arrivés. Le temps nécessaire à l'envoi du paquet puis de l'accusé de réception correspond au débit de la connexion et non au débit du multiplex.

Par exemple, si l'application se voit offrir une connexion à 10 Mbit/s, le contrôleur espaceur va émettre les cellules du paquet TCP au rythme de 23584 cellules/s, soit un paquet toutes les 1,35 ms. Les accusés de réception seront émis au même rythme. Il n'est pas nécessaire d'espacer les accusés de réception car ils tiennent dans seulement 1 à 2 cellules. La fenêtre de congestion sera donc augmentée au rythme de la taille d'un paquet toutes les 1,35 ms soit au débit de 10 Mbit/s. TCP est donc régulé non seulement au niveau ATM par l'émission des cellules, mais surtout en accordant des crédits supplémentaires à la fenêtre de congestion. Le même résultat peut être obtenu si le flux TCP est envoyé sur une liaison synchrone à débit faible e.g. 2 Mbit/s. La régulation est ici produite par la technologie de transfert de la couche physique. La seule contrainte du dispositif est de devoir stocker une fenêtre d'émission complète, soit de 8k à 64k octets⁹, sans perte.

De façon similaire [97], l'espacement des accusés de réception produit le même effet sur la régulation de la taille de la fenêtre de congestion et donc sur la régulation du débit. Des essais pour modifier à la volée, i.e. par un équipement du réseau, la valeur de taille de la fenêtre en réception (qui implique celle d'émission) conduisent également au même type de

⁹ La taille de la fenêtre d'émission de TCP dépend du réglage de la pile TCP et du système d'exploitation. Les valeurs courantes étant de 8k, 16k, 32k et 64k octets.

régulations. Cependant, les cellules continuent à être émises au débit du multiplex sous forme de rafales de 32 cellules, ce qui peut occasionner des pertes dans les files d'attente des commutateurs à la différence de l'espacement qui les protège des rafales. Une implémentation différente [98] basé sur l'insertion de paquets vide par le terminal lui-même permet également d'espacer l'émission des paquets TCP directement à la source.

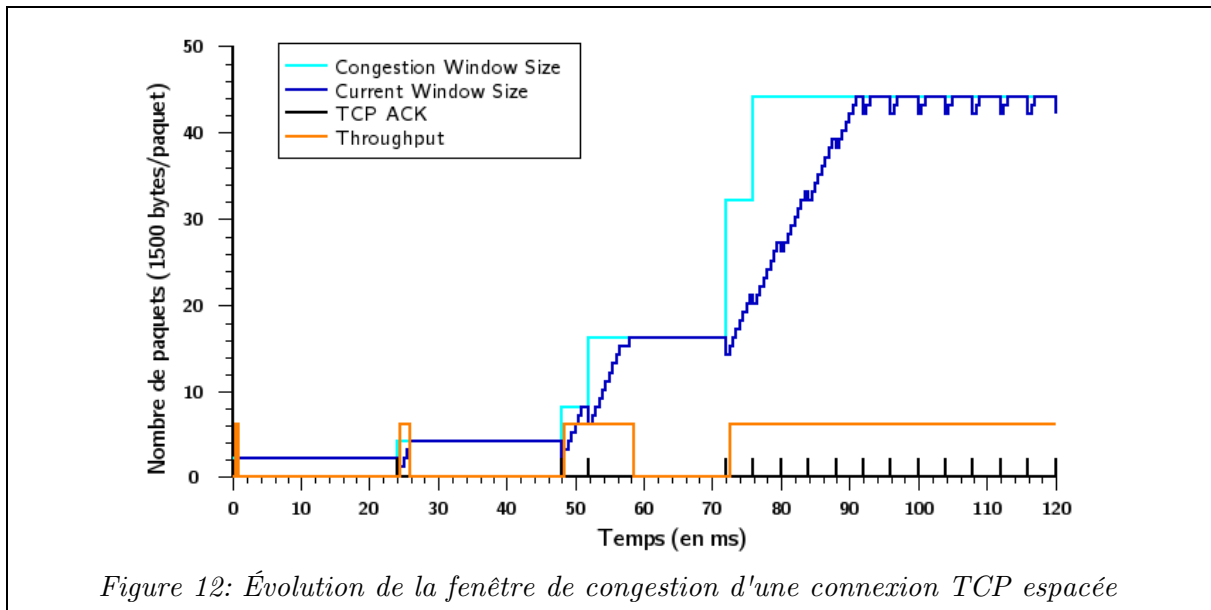
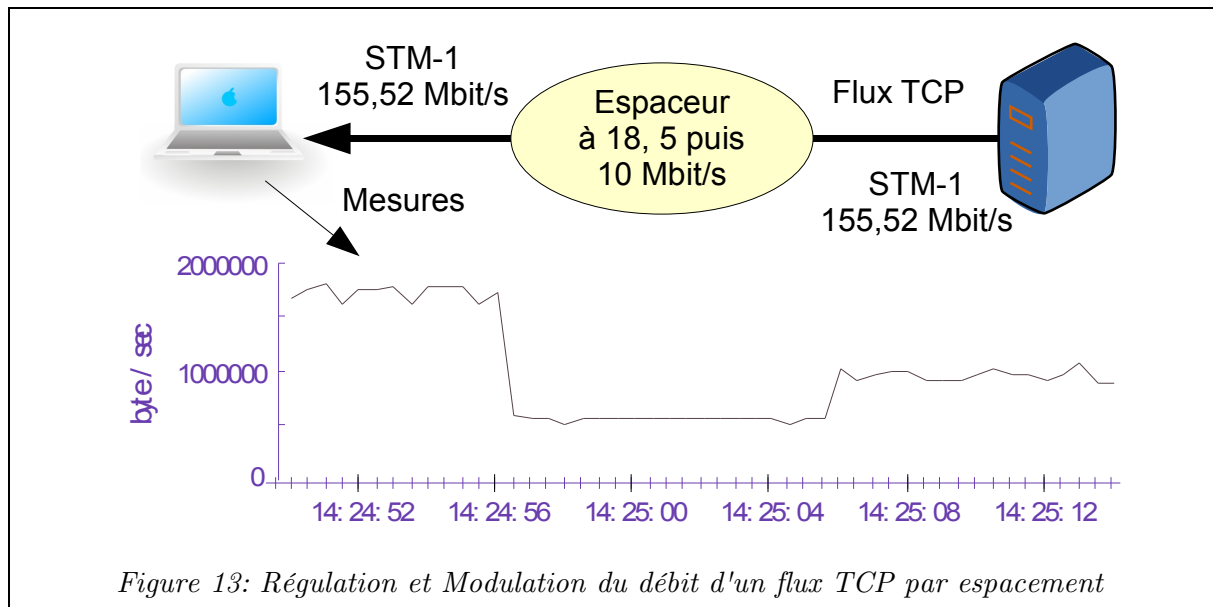


Figure 12: Évolution de la fenêtre de congestion d'une connexion TCP espacée

Finalement, il est possible de moduler le débit d'espacement sans occasionner de perte au niveau de la connexion TCP (voir figure 13). Si le débit est réduit, l'émission des paquets suit puisque la taille de la fenêtre est augmentée au rythme de ce nouveau débit. De même, en cas d'augmentation du débit, la connexion TCP profite de ce surcroît de débit, car la taille de la fenêtre augmente plus rapidement permettant à TCP d'envoyer plus de paquets par unité de temps. Par analogie, il fut observé et démontré que si n connexions TCP sont espacées globalement i.e. au niveau du VP au lieu du VC, chaque connexion va automatiquement s'adjuger $1/n$ de la bande passante du VP. Il y a partage équitable du débit grâce à nouveau au mécanisme de contrôle des émissions par fenêtre. À l'inverse, UDP n'ayant pas de contrôle, il n'est pas possible de réguler le débit UDP sur une valeur donnée. Seul l'espacement au débit nominal de l'application est possible. De plus, si on mélange des flux UDP et TCP, les connexions TCP s'écrasent devant les connexions UDP. Les flux TCP ne se partagent plus $1/n$ du débit, mais $1/n$ du débit laissé par les connexions UDP, qui peut s'avérer être proche de 0.

L'ITU-T a également généralisé l'algorithme du GCRA aux paquets à longueur variable comme l'IP. Appelé « Generic Byte Rate Algorithm - GBRA » [99], il opère sur une quantité donnée d'octets e.g. 40, en place et lieu des cellules ATM. Sa mise en œuvre est cependant plus délicate. Lorsqu'un paquet arrive, il faut tout d'abord extraire la taille de ce paquet dans l'en-tête IP avant de pouvoir exécuter le GBRA. Ensuite, l'espacement réel doit tenir compte de la taille du paquet pour l'affectation du média de sortie puisque tous les paquets n'ont pas la même taille. Si cela est quasi transparent pour les petits paquets lorsque le débit physique de l'interface est élevé, il peut s'avérer inefficace sur des liaisons à faible débit. Par exemple, sur une connexion Wifi avec un débit de 1 Mbit/s seulement, l'émission d'un paquet de données de 1500 octets prend 12 ms. Si durant cet intervalle, un petit paquet

d'une application de VoIP doit être émis, il subira une gigue maximale de 12 ms qu'il n'aurait pas eu à subir si la couche ATM était présente. La mise en cellule de l'ATM permet de s'affranchir de ce problème car elle autorise l'entrelacement des paquets ce qui n'est pas possible en couche IP.



1.4. Conclusion

Les principes de l'espacement et de la régulation de TCP ont été présentés lors de divers salons et conférences. Le VSA, en tant que norme, et l'espacement, en tant que principe de gestion de la QoS, ont été implémentés par de nombreux constructeurs d'équipements dans leurs produits ATM.

La première contribution majeure de ce mémoire est de proposer un mécanisme de gestion des files d'attente et de multiplexage permettant de réguler le débit des connexions TCP sans perte à condition de respecter les règles suivantes :

Règle n°1 : L'équipement réalisant l'espacement au débit le plus faible doit pouvoir stocker sans perte toutes les cellules correspondant à la taille d'une fenêtre d'émission TCP, pour chaque connexion espacée.

Le nombre de connexions TCP pouvant bénéficier d'une QoS garantie dépend de la capacité de l'équipement. Un premier niveau d'admission d'appel doit donc être effectué pour respecter cette règle.

Règle n°2 : Les connexions TCP doivent être espacées indépendamment des connexions UDP, en affectant des VC ou VP dédiés aux connexions TCP et d'autres pour l'UDP.

Le débit et la QoS seront garantis au protocole TCP à condition qu'il ne soit pas en présence de connexions UDP.

2. La signalisation allégée

2.1. Présentation du projet ASIA

Le projet RNRT ASIA « Accelerated Signalling of Internet over ATM » [100] s'est déroulé dans un contexte de recherche où les efforts étaient portés sur la meilleure technique pour transporter les flux Internet (TCP et UDP) tout en garantissant de la QoS. Plusieurs propositions de signalisation dite allégée, par opposition aux limitations de RSVP, furent élaborées à la même époque : UNITE [101], OPENET [102], et Dynaflo [103]. L'architecture ASIA permet de partager des ressources pré-établies par MPLS-TE à l'aide des cellules de réservations (RM) des ATC ABT/DT et ABT/IT. L'originalité de la solution repose sur le couplage des messages TCP avec l'émission des cellules RM et dans l'application de la régulation de TCP par espacement dont le débit est modulé par les cellules RM qui réservent les ressources disponibles sur le VC.

Les contributions de l'auteur au projet ASIA portent sur :

- l'architecture en trois plans : service , contrôle et transfert,
- l'usage de l'ATC ABT/IT pour moduler le débit des connexions TCP,
- l'approvisionnement des connexions ATM ABT à l'aide de MPLS-TE,
- la synchronisation de TCP avec la réservation de débit par cellule ABT/DT.

2.2. Couplage ABT et MPLS

En premier lieu, il est nécessaire de disposer d'une connexion ATM établie au niveau du VC pour pouvoir utiliser la réservation rapide de débit à l'aide de cellules RM. Comme le but du projet ASIA était de coupler l'ATM et l'IP [104], c'est vers les protocoles de l'IETF et non ceux de l'ITU-T que se sont orientées les études. MPLS-TE, dans le cadre de la solution ASIA, est utilisé pour établir les connexions ATM entre les nœuds (LER) d'entrées et de sorties du réseau [105].

La variante « Traffic Engineering » de MPLS permet de réserver des ressources pour les tunnels portés par des VC ATM. Un premier niveau d'admission d'appel est exécuté par chaque équipement MPLS (LER ou LSR) afin de vérifier si la connexion peut être établie. Ces VC (tunnels), sont établis avec un débit minimum, e.g. 1 Mbit/s, et maximum e.g. 10 Mbit/s conformément à l'ATC ABT/IT. Pour chaque interface, une réserve de bande passante est également attribuée à l'ensemble des connexions de capacité de transfert ABT/IT. Par défaut, les tunnels MPLS-TE sont déclinés en tant que VC UBR dans un réseau ATM. Cette réserve de bande passante est également prise en compte par la CAC lors de l'établissement des tunnels de sorte à ce que cette réserve reste suffisante pour un nombre donné de connexions.

2.3. Partage des ressources avec l'ABT/IT

Lorsqu'un service utilise un tunnel MPLS-TE approvisionné par le système ASIA, il peut bénéficier d'un surcroît de bande passante au fur et à mesure de son besoin. Pour cela, le nœud d'entrée du réseau qui réalise la fonction de contrôle et d'espacement au rythme du débit initial du VC, surveille en permanence l'état de la mémoire associée à cette connexion. Si le nombre de cellules dépasse un seuil prédéfini, des cellules RM ABT/IT sont émises à intervalles réguliers e.g. 100 ms, sur cette connexion pour demander plus de débit i.e. le débit maximum du VC, jusqu'à ce qu'un seuil inférieur soit franchi dénotant une faible occupation de la mémoire. Tous les nœuds du réseau i.e. LSR, vont traiter la cellule RM selon

l'algorithme de partage de la bande passante suivant :

- La première cellule RM permet de notifier les serveurs RM i.e. les LSR et le dernier LER (egress node), qu'une nouvelle connexion demande du débit supplémentaire. Si il reste du débit de réserve, la cellule RM poursuit son chemin. Si il reste plus ou moins de débit que celui demandé, la cellule RM est modifiée en conséquence. Au final, le LER demandeur reçoit via la cellule RM le nouveau débit alloué à la connexion et ajuste en conséquence le débit de contrôle et d'espacement. Ce débit peut éventuellement correspondre au débit nominal de la connexion si il ne reste plus du tout de débit de réserve,
- Les serveurs RM vont alors allouer $1/(n+1)$ au lieu de $1/n$ du débit de réserve à chaque connexion au fur et à mesure de la réception des cellules RM ABT/IT correspondantes de sorte à laisser $1/(n+1)$ du débit de réserve disponible pour la $n+1$ connexion
- La deuxième cellule RM ABT/IT de la $n+1$ connexion se verra attribuer un débit égal aux autres connexions.
- Lorsque la mémoire du contrôleur espaceur est vide ou passe en dessous d'un seuil, une dernière cellule RM ABT/IT, avec un débit nul, est envoyée pour libérer la bande passante qui est remise dans le pool de débit de réserve.

Cet algorithme permet de partager équitablement le débit entre n connexions tout en garantissant un débit minimum correspondant au débit initial de la connexion. Cet algorithme permet d'allouer du débit de la même façon que l'ATC ABR de l'ATM Forum, mais avec une garantie de débit et de QoS équivalente à une ATC DBR.

2.4. Synchronisation de TCP avec ABT/IT

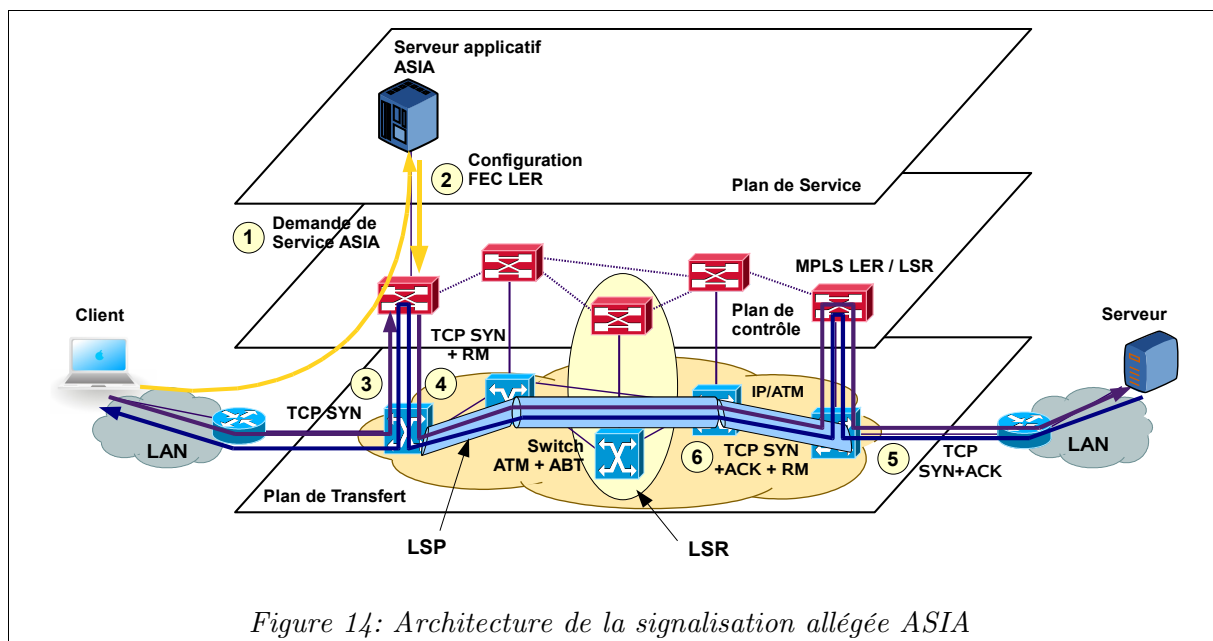
Pour compléter le dispositif, un mécanisme [106], présent dans les nœuds d'entrée du réseau, permet de détecter les débuts et fins des connexions TCP afin d'initialiser le contexte afférent et leur attribuer le bon VC. En effet, la technologie MPLS est associée avec la notion de classe d'équivalence de routage « Forwarding Equivalent Class - FEC » sur les LER pour déterminer quel tunnel utiliser en fonction de l'en-tête des paquets TCP/IP. L'originalité du projet ASIA est de compléter les FEC à la volée.

Une application désirant utiliser le système ASIA (voir figure 14) effectue une demande de service (1) auprès de la plate-forme de supervision en précisant simplement son adresse source et l'adresse destinataire. La logique de service se charge alors de déterminer quel est le LER auquel l'application est raccordée, puis configure (2) la FEC correspondante pour cette application dans le LER. Ensuite tout est automatisé.

Lorsque le LER reçoit le premier message TCP SYN (3), il analyse l'en-tête du paquet TCP et en extrait les informations de ports et d'adresses source et destination pour compléter la FEC. Ceci permet au LER de déterminer quel tunnel MPLS-TE la connexion doit emprunter. Puis, il émet une cellule RM ABT/DT (4) sur cette connexion afin de réserver le débit minimum garanti. Si ce débit minimum ne peut être garanti, un ou plusieurs serveurs ABT, localisés dans les LSR, vont rejeter l'appel et retourner une cellule RM ABT/DT portant l'indicateur d'échec (NACK). La FEC initialisée précédemment est détruite ainsi que le message TCP SYN. L'application essaye de nouveau d'établir la connexion conformément au protocole TCP. La réservation peut alors aboutir. Si elle n'aboutit pas, l'application s'arrête sur un constat d'échec et le réseau est protégé contre une surcharge (fonction de CAC). En cas de succès, le message TCP SYN est envoyé dans le tunnel au destinataire. Le LER de sortie, averti par la cellule RM, va également créer une FEC pour la demi connexion

retour. Le serveur envoie à son tour un message TCP SYN + ACK (5) qui va être traité de la même façon (6), mais sur le sens serveur vers client. Une fois la poignée de main du protocole TCP terminée, les LER de chaque extrémité sont configurés avec la FEC correspondant à cette nouvelle connexion TCP. L'algorithme ABT/IT de partage de débit peut être déroulé.

En cas de perte de cellule RM ou de refus de réservation de débit sur le demi sens serveur client, le LER d'entrée ne réalise aucune opération lorsqu'il voit passer la deuxième tentative d'établissement de la connexion TCP. Une temporisation égale à la temporisation maximale de TCP est également armée sur le LER d'entrée lorsqu'il voit passer le premier message TCP SYN. Cette temporisation permet de libérer la FEC et le contexte afférent en cas de refus sur l'autre demi sens de la connexion ou en cas de perte de paquet de sorte à ne pas saturer la table de connexions ni consommer inutilement du débit.



Finalement, les LER pistent les messages TCP FIN de sorte à libérer le contexte, la FEC et le débit minimum réservé. Chaque libération est effectuée indépendamment l'une de l'autre comme dans le protocole TCP. Seule la connectivité reste sans avoir la possibilité de réserver du débit, ce qui est inutile puisque cette direction ne transmet que les messages TCP ACK. Pour pallier au problème des demi-connexions TCP qui ne se ferment¹⁰ pas, une temporisation longue est armée lorsque la mémoire d'espacement est vide. Afin de ne pas mobiliser des ressources trop longtemps, le contexte, la FEC et le débit minimum sont relâchés à l'expiration de la temporisation.

2.5. Conclusion et bilan

L'apport du projet ASIA à la compréhension du comportement de TCP en présence du contrôle espacement a permis d'édicter les règles énoncées précédemment. Mais l'apport le plus significatif réside dans l'architecture même du projet. Il a permis de poser les bases de la

10 Cas de l'attente de TCP dans l'état FIN_WAIT_STATE dû aux applications comme les navigateurs web en HTTP/1.0 qui ne ferment jamais leur demi connexion

solution globale énoncée en fin de partie. D'une part, il est nécessaire d'approvisionner le réseau à l'aide de tunnels MPLS-TE qui sont ensuite protégés par une fonction d'admission d'appel. D'autre part, le couplage entre application et réseau est nécessaire pour s'affranchir du couplage trop fin entre protocole de transport et le réseau. Ainsi, les principes élaborés pour le protocole TCP ne sont pas transposables au protocole UDP. Il faut remonter dans les couches pour de nouveau pouvoir synchroniser, par exemple, RTP avec le réseau. Enfin, la déclaration préalable de l'application dans le service ASIA et la recherche du nœud de raccordement, sont les prémices du besoin de découplage service/contrôle/réseau et de la connaissance de la topologie du réseau. Les principes d'ASIA ont été implémentés par Ericsson, partenaire du projet, sur leur commutateur ATM et présentés aux salons « ATM Development ».

La disgrâce de la technologie ATM, rattrapée par l'IP et les études sur le NGN, ont eu raison de la signalisation allégée. ASIA ne trouva jamais de débouché, mais a constitué un formidable « melting pot » sur lequel le reste des études de ce mémoire furent élaborées.

3. La médiation d'accès aux ressources

3.1. Présentation du projet Cadenus

Le projet CADENUS [107] avait pour objectif de développer une architecture permettant à un utilisateur de choisir un service avec une Qualité de Service prédéfinie ou adaptée à ses besoins. L'architecture est basée sur la notion de **composant de médiation**. Chacun joue un rôle particulier qui est essentiel pour l'admission d'appel, la traduction de contrat et la commande de ressources. L'idée principale est de traduire un contrat de service « Service Level Agreement - SLA » en une ou plusieurs spécifications de service « Service Level Specification - SLS ». A partir de cette spécification, un médiateur de ressources ainsi qu'un contrôleur de réseaux exécutent finalement la commande de ressources. L'architecture a été conçue comme l'application de l'outil de classification OURAf. La décomposition fonctionnelle a été implémentée organiquement selon plusieurs plates-formes de services. Les contributions [108] de l'auteur au projet CADEDNUS ont porté essentiellement sur :

- La définition de l'architecture,
- La définition du modèle de rôle et d'acteur au sein de l'architecture,
- La définition de l'architecture et la réalisation à partir de produits commerciaux du Resource Manager.

3.2. Architecture basée sur la médiation

L'architecture choisie pour CADENUS (voir figure 15) se compose de 6 blocs : l'Access Mediator (AM), le Service Mediator (SM), les Services de contenus, le Service Authority and Directory, le Resource Mediator (RM) et le Network Controller (NC) [109].

- **L'Access Mediator (AM)** ou Médiation d'Accès aux Services (MAS) offre un accès unique aux fournisseurs de services pour les clients. Elle assure l'authentification unique de l'utilisateur, l'agrégation de toutes les offres de services et leur présentation unifiée.
- **Le Service Mediator (SM)** ou Médiation de service (MS) fournit la présentation du service, des abonnements, des contrats, du profil utilisateur et l'accès au service de contenu. C'est le point d'entrée unique pour les services de contenus.
- **Le service de contenus** exécute la logique de service. La logique de service correspond aux données requises pour effectuer le service. Par exemple les serveurs Vidéo.
- **Le Service Authority and Directory** ou autorité et annuaire de service stocke les descriptions de service et fournit une référence du niveau de la qualité de service. Il assure une homogénéité des niveaux de QoS pour l'ensemble des services qu'il répertorie.
- **Le Resource Mediator (RM)** ou Médiation d'Accès aux Ressources (MAR) est exigé en raison de la multiplicité des opérateurs de réseau, de la multiplicité des technologies impliquées dans la couche de commande de ressources, et des différents niveaux de QoS qui peuvent être offerts. Il a la connaissance de bout en bout de la QoS, identifie les différentes partitions de réseau impliquées dans la QoS demandée et effectue l'adaptation entre le service et la commande de ressources. Il choisit les ressources appropriées à mettre en œuvre pour satisfaire les demandes de service et fournit une vue abstraite du réseau au service de contenu mais ne commande pas directement les ressources.

- **Le Network Controller (NC)** ou contrôleur de réseau (CR) commande les ressources à l'intérieur de son domaine et exécute la configuration des équipements du réseau. Son rôle principal est de traduire un ensemble de règles indépendantes de la technologie, fournies par l'RM, en un ensemble de règles dépendantes de la politique du réseau et de la technologie sous-jacente.

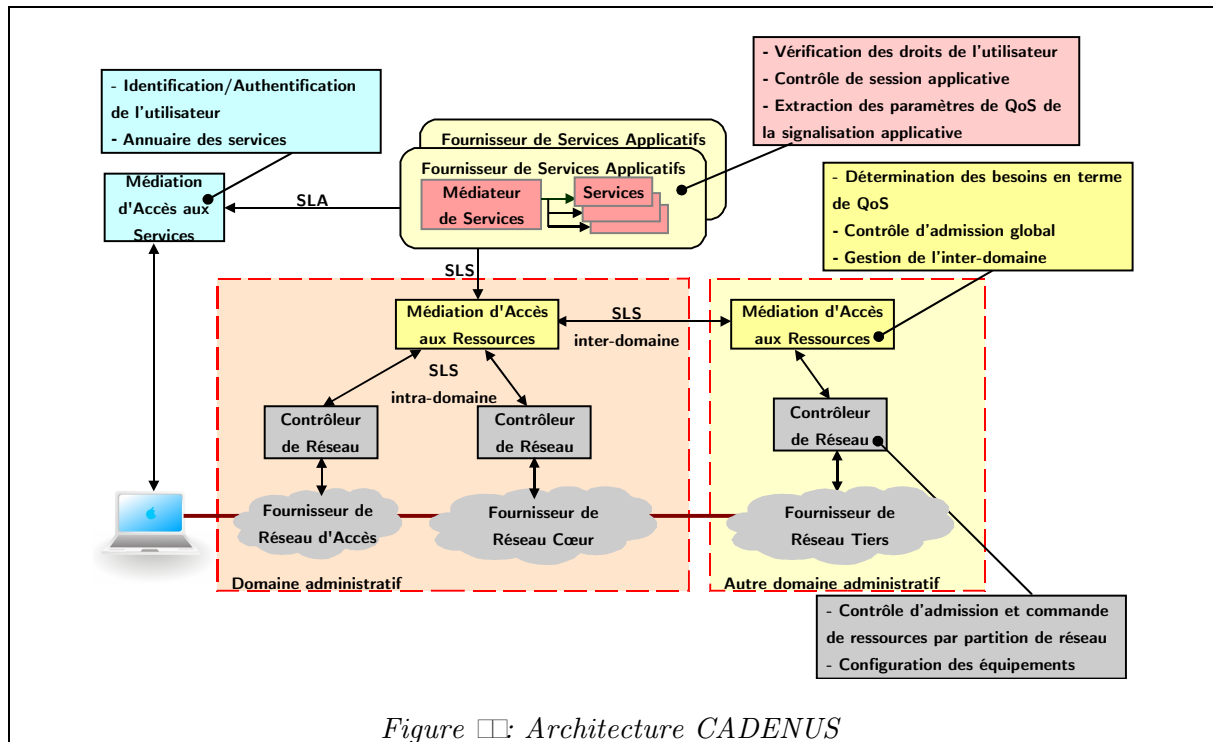


Figure 11: Architecture CADENUS

L'architecture CADENUS est régie par deux idées principales : la médiation généralisée pour tous les composants, et la négociation et la traduction de contrat entre tous les composants. L'échange de SLA et de SLS peut être mis en application de différentes manières, comme avec COPS par exemple [110]. Cependant, les fonctions principales doivent être celles d'un système réparti : transactionnel, fiable, performant et sécurisé. L'eb-XML et les « Web-Services », deux des systèmes les plus connus de l'Internet, ont été conçus dans l'optique du commerce électronique (e-business). Le processus de négociation de la QoS étant semblable à un système marchand, le concept d'eb-XML a été choisi comme vecteur de transport pour l'implémentation de l'architecture [111].

3.3. Du SLA au SLS

À l'origine, il existe deux types de contrat SLA :

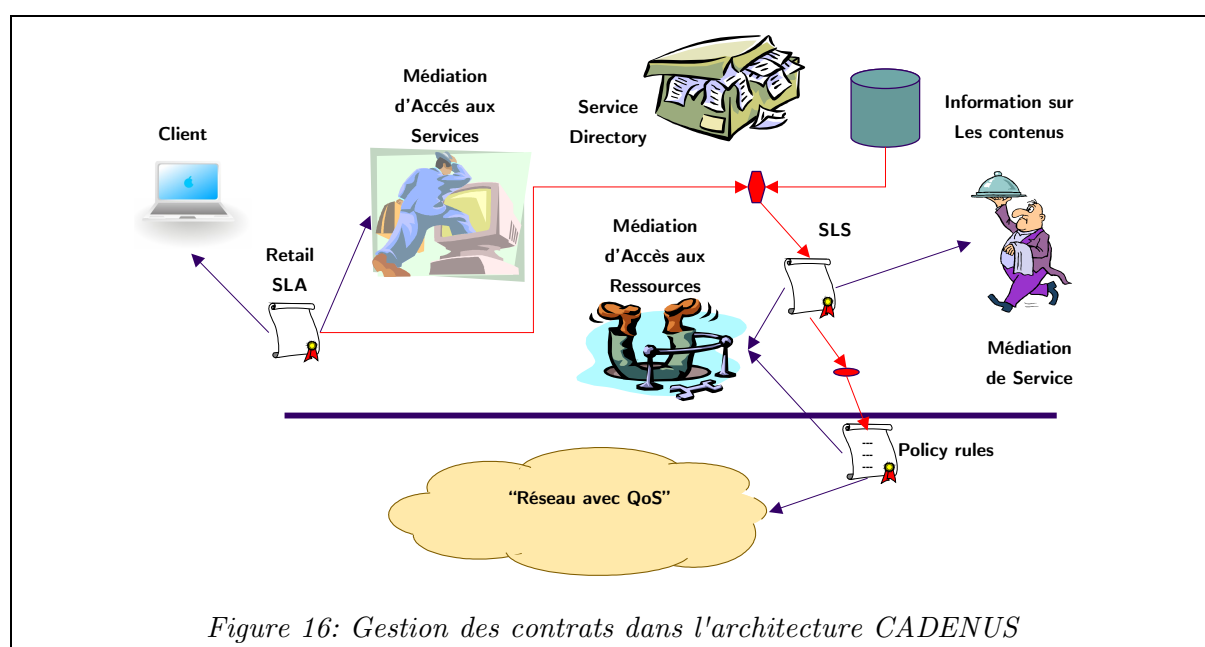
- **le retail-SLA (r-SLA)** pour la vente au détail entre le client et le service unitaire,
- **le wholesale-SLA (w-SLA)** pour la vente en gros entre fournisseur de service et fournisseur de réseau.

L'utilisation traditionnelle du SLA (voir figure 16) a lieu dans le processus de commande/livraison et il est employé quand un client souscrit ou modifie son contrat à un service. Le niveau global de QoS auquel le client peut prétendre peut alors être négocié. Le SLS correspondant à ce SLA, sert principalement à la planification des ressources du réseau. Ces deux étapes sont traitées en différé sans contrainte temps réel.

Même si ce type de contrat est bien conçu pour l'abonnement et l'approvisionnement, il n'est pas approprié au processus d'invocation : il n'est pas utile de modifier son contrat lorsque l'on passe un appel téléphonique. À cette fin, un nouveau contrat appelé **SLA d'invocation (i-SLA)** est défini. L'**i-SLA** est exclusivement employé pendant le processus d'utilisation. Par définition, il est traité en ligne avec une forte contrainte temps réel. D'une manière similaire, un **SLS d'invocation (i-SLS)** est défini pour la commande de la QoS. L'**i-SLA** est gardé aussi simple que possible. Il décrit juste le niveau de QoS que le client demande. Le **i-SLA** ne contient que les informations nécessaires à l'invocation du service en faisant référence au SLA sans toutefois être redondant avec celui-ci. L'**i-SLS** est plus proche du SLS normal. Il contient tous les paramètres nécessaires à l'opérateur de réseau pour commander les ressources.

Tout d'abord, la médiation de service reçoit de la médiation d'accès un **i-SLA**, qu'elle va devoir traduire en **i-SLS** selon différents scénarios. Elle récupère dans une base de données un **i-SLS** correspondant au service décrit dans le **i-SLA** ou demande au service unitaire de lui fournir un **i-SLS** pré configuré qu'elle n'aura plus qu'à compléter. Elle peut être également court-circuitée, auquel cas, c'est le service unitaire qui est directement impliqué dans la production de l'**i-SLS**.

Enfin, quand la médiation de service agit en tant que Proxy de signalisation, l'**i-SLA** est présent dans le protocole e.g. la liste des codecs dans SIP. L'**i-SLS** est alors directement calculé à l'aide des paramètres contenus dans le message de signalisation e.g. le codec utilisé déterminera le débit nécessaire. Le **i-SLS** est ensuite transmis au RM avec lequel le marché a été conclu. Il affine le niveau technique du **i-SLS** et transmet autant de SLS que nécessaire vers les NC concernés. Une fois la configuration en place, le service peut être délivré à l'utilisateur. Lorsqu'elle est impliquée, la médiation de service sert une dernière fois d'intermédiaire pour demander au service unitaire de délivrer le contenu.

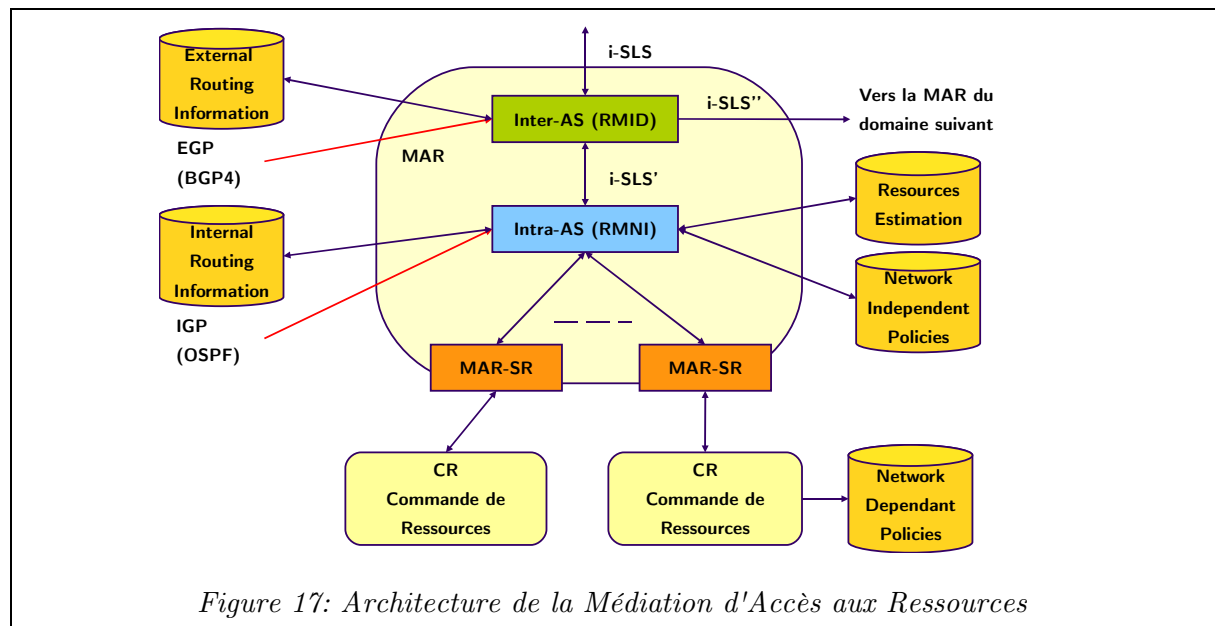


Dans le cas où plusieurs opérateurs de réseau sont impliqués, les RM respectifs doivent pouvoir s'échanger des demandes d'établissement de QoS. Un premier accord cadre, de gré à gré, est établi entre les RM afin qu'ils puissent s'échanger des **i-SLS**, ceci afin de ne pas

renégocier la QoS pour chaque appel. Les opérateurs conviennent d'un volume de QoS qui peut être échangé sans renégociation. Ce volume de QoS peut varier dans le temps en fonction des besoins respectifs de chacun. Le i-SLS échangé entre les RM est identique (aux valeurs de paramètres près) à celui échangé entre la médiation de services et de ressources.

3.4. Architecture du Resource Mediator

Les fonctions du RM (voir figure 17) sont la localisation des terminaux et des serveurs, le choix des réseaux et/ou la partition de réseaux, l'agrégation de l'information et l'admission d'appel de bout en bout. Il se décompose en trois parties et ne gère que la couche indépendante de la technologie de transport.



3.4.1. Fonction Inter Domaine

Cette fonction a la vue globale de tous les domaines et de bout à bout. Elle correspond au point d'entrée unique pour chaque fournisseur de réseau et reçoit toutes les demandes venant des autres médiations (de service ou de ressources). Sa première tâche est de calculer, à partir de l'adresse IP destination et/ou source, la portée du i-SLS. Il détermine ainsi si la QoS à mettre en place concerne uniquement son domaine ou si d'autres RM doivent être impliqués. Il utilise une base de données topologiques alimentée à l'aide d'informations issues des protocoles de routage externe, comme BGP4 et/ou par des informations de localisation géographique.

Dans le cas d'un i-SLS multi domaine, le RM découpe l'i-SLS en un i-SLS' qui concerne son propre domaine et un i-SLS'' pour les domaines suivants. Ces nouveaux i-SLS correspondent à l'i-SLS original augmenté des points d'interconnexion des deux réseaux. Ils sont conformes avec la description traditionnelle du SLS présentée dans [5]. L'i-SLS'' est envoyé au RM du domaine suivant et l'i-SLS' est envoyé à la fonction intra domaine. Le premier niveau d'admission d'appel est exécuté à ce stade. Il s'agit simplement de vérifier la conformité du i-SLS. En retour, les réponses, en provenance des autres RM ou des NC de son domaine, seront agrégées afin de fournir une réponse unique à la médiation de service. Cette fonction a la responsabilité de la QoS de bout en bout. La gestion du temps (début, fin

réserve à l'avance) est du ressort de cette fonction. Dans le cas des réservations en avance, selon un calendrier, elle déclenche la mise en place automatiquement lorsque la demande arrive à échéance.

3.4.2. Fonction Intra Domaine

Version intra domaine de la fonction précédente, elle ne gère que son domaine et choisit le mode de transfert le plus approprié pour chaque partition des réseaux impliqués. L'i-SLS' est à nouveau découpé pour chaque partition de réseau. Ce travail est effectué à partir de la base de topologie alimentée par les protocoles de routage interne (OSPF ou IS-IS) ou toute autre information de localisation. Un deuxième niveau d'admission d'appel basé sur une estimation globale des ressources permet de rejeter un i-SLS' en cas de surcharge des réseaux impliqués. Cette fonction a la responsabilité de la QoS dans son domaine.

3.4.3. Adaptation aux CR

Cette dernière partie a seulement la vue de la partition de réseau qu'elle contrôle. Elle adapte la spécification i-SLS reçue aux contraintes technologiques du mode de transfert sous-jacent. Dans le cas du niveau 2, cette fonction doit récupérer les informations nécessaires à la commande de ressources avant de transmettre la demande au NC. Il est de la responsabilité du RM de déterminer et informer le NC quand un service réseau doit être activé, arrêté, annulé ou modifié. Le NC est sous le contrôle total du RM.

3.5. Le Contrôleur de Réseaux

Si le RM a une vue de bout en bout, la portée du contrôleur de réseau est limitée au réseau (un seul domaine) ou à la partition de réseau (Accès ADSL, ATM, ...) dont il est responsable. Il est totalement dépendant de la technologie et peut supporter de une à plusieurs technologies de transfert. L'architecture impose donc de déployer au maximum un NC par technologie de transport supportée. En plus de la commande des équipements, l'admission d'appel s'effectue sur les données locales qu'il a pu collecter auprès des équipements du réseau. Le NC peut être amené à traduire des informations concernant la couche de transport 3 (IP) en informations de couche 2 (ATM, VLAN, ...). Les données nécessaires à cette translation peuvent être fournies par le RM ou récupérées par le NC auprès d'autres entités (serveur d'authentification RADIUS pour l'ADSL par exemple). Le NC offre un niveau d'abstraction pour la mise en place de la QoS afin d'être le plus indépendant possible des équipements et de leur propre système de commande (CLI, SNMP, COPS, XML, ...). Le NC peut également configurer les équipements en utilisant une signalisation (ie. PNNI pour l'ATM, CR-LDP/RSVP-TE pour MPLS). Dans ce cas, il agit en tant que Proxy de signalisation.

3.6. Modèle « Hub » versus « Cascade »

Deux philosophies pour la gestion de la QoS de bout en bout ont été étudiées dans le projet CADENUS. Apparues dans [112], elles distinguent deux modèles selon le comportement de la couche traitement de service par rapport à la couche de traitement des ressources.

Le modèle « Hub », en référence au principe de connexion Ethernet, laisse la responsabilité de la garantie de QoS de bout en bout à la médiation de service. Cela nécessite de contacter tous les RM en charge des réseaux impliqués dans la connexion. La médiation de service doit donc non seulement connaître les différents réseaux empruntés par la connexion, mais également leur interconnexion. Si ce modèle est préféré des fournisseurs de services car

ils maîtrisent toute la chaîne, il n'est pas apprécié des opérateurs. En effet, ils ne maîtrisent plus complètement leurs relations entre eux, ni le chemin inter domaine emprunté par la connexion. La relation entre service et contrôle peut être enrichie pour que le RM retourne le point d'interconnexion qu'il a choisi avec l'opérateur suivant. Cependant, le choix du chemin d'AS est dicté par la médiation de service. De plus, la nécessaire connaissance et le choix du chemin d'AS par le SM ne respectent pas le principe de séparation des couches. Ce modèle est partiellement appliqué dans le cadre de l'IMS. Le routage de bout en bout est assuré au niveau du protocole SIP indépendamment du routage au niveau transport. L'alternative réside dans la possibilité pour les opérateurs d'effectuer une configuration cohérente des S-CSCF et des routeurs puisqu'ils maîtrisent l'ensemble : cœur IMS et réseau IP.

A l'inverse, le modèle « Cascade » laisse la responsabilité de la garantie de QoS au niveau du RM. La médiation de service s'adresse au RM du réseau où elle est connectée. Préféré des opérateurs de réseau, ce modèle laisse toute la latitude pour choisir non seulement le meilleur chemin d'AS mais également les points d'interconnexions entre les réseaux de transports. La requête de QoS est propagée de proche en proche en suivant le chemin d'AS choisi par le premier RM. Ce dernier est responsable de la garantie de la QoS de bout en bout. En cas de problème, il doit interroger ou récupérer des autres RM toutes les informations (alarmes, mesures, ...) nécessaires à la localisation et la correction de la mise en place de la QoS. L'indépendance entre service et contrôle est bien respectée et la médiation de service n'a pas besoin de connaître les détails des réseaux des opérateurs [113].

3.7. Conclusion et Bilan

L'application de l'outil OURAf pour construire l'architecture CADENUS a permis de dégager plusieurs fondements sur la gestion de QoS. En premier lieu, il est impératif de segmenter la mise en place de la QoS entre une couche indépendante de la technologie, qui possède une vision de bout en bout, et une couche adaptée aux impératifs de la technologie de transport sous-jacente. Une deuxième découpe en fonction des partitions rend le modèle extensible et applicable aussi bien au réseau de grande taille qu'à l'inter-domaine. La création et la gestion de services est elle aussi bien segmentée et le modèle ebXML se conforme bien aux recommandations du projet [114]. Si plusieurs démonstrations des principes de CADENUS ont validé le concept, tous les problèmes soulevés n'ont cependant pas trouvé de réponse. En particulier, la mise en place de la QoS au niveau des processus de commande/livraison et la surveillance des SLA/SLS n'ont pas été abordés. Les fondements de l'architecture ont cependant servi de guide pour le projet EuQoS qui a permis de trouver des solutions aux problèmes en attente tels que la continuité de QoS, la déclinaison de l'architecture par technologie, l'approvisionnement et la surveillance des ressources et la mise en oeuvre de la négociation de QoS à l'inter-domaine. Les principes de CADENUS ont été réalisés à l'aide d'équipements commerciaux et présentés lors de la conférence « World Telecommunication Congress » à Paris ainsi qu'à l'état major de France Telecom.

Les contributions de ce mémoire au projet CADENUS ont ainsi pu conforter les prémisses des problèmes apparus avec l'architecture de signalisation allégée ASIA durant cette période. Si l'acquisition de la topologie pouvait trouver une solution au travers de l'écoute des protocoles de routage, elle ne répondait que partiellement à la problématique et limitait fortement le champ d'application aux réseaux d'opérateurs. Le besoin de normaliser une interface entre service et contrôle s'est précisé. Le modèle de commande centralisé n'était pas viable du fait des performances de la configuration des équipements nettement inférieures à ce qui est nécessaire dans un réseau de grande taille. Seule la fonction d'admission d'appel pouvait se généraliser mais nécessitait de travailler de nouveau la partie approvisionnement afin de pallier aux problèmes de configuration dynamique du réseau.

4. Service de QoS à la demande

Disposer de mécanismes de gestion de la QoS dans le réseau est une chose. Il faut impérativement donner la possibilité aux applications non seulement d'exprimer leur besoin de QoS, mais surtout de pouvoir déterminer avec précision quand (début et fin) et où (source et destination) cette QoS doit s'appliquer. Il existe pléthore de classifications des CoS, mais aucune API n'est réellement disponible ou a rencontré le succès escompté pour que les applications s'interfacent avec le réseau. Seuls les mécanismes à base de signalisations offrent cette possibilité. C'est dans ce contexte, suite aux travaux dans le projet CADENUS, que fut élaboré un service de QoS à la demande pour négocier les SLS. Si l'interface, entre la médiation de ressources et de services, est bien identifiée, elle n'était pas standardisée. Dans certains cas, elle était inaccessible de l'extérieur lorsque la plate-forme de service embarque une logique de gestion des ressources. Afin de rendre la commande de ressources multiservice, il est impératif de disposer au-dessus de la médiation d'accès aux ressources, d'un point d'accès au service « Service Access Point -SAP ». La contribution de l'auteur a consisté :

- À identifier l'ensemble des paramètres de QoS pour toutes les applications, puis de rendre optionnels ceux qui ne servent pas à tous les services ou qui ne sont pas strictement nécessaires à la mise en place de la QoS,
- À définir les fonctions afférentes à la manipulation de ces paramètres, et les contraintes et propriétés que doit respecter le protocole de transport de ces paramètres et fonctions,
- À implémenter plusieurs déclinaisons de cette interface (de la simple socket TCP jusqu'au web services).

4.1. Description

Le service de QoS à la demande, appelé Qu@ds, est un service réseau accessible par n'importe quelle logique de service, qu'elle soit de type proxy (SIP, RTSP), serveur WEB ou plate-forme de service. Il normalise l'accès à la fonction d'admission d'appel et de réservation de ressources qui contrôle le réseau.

4.1.1. Paramètres

Les paramètres, nécessaires à l'établissement du contrat SLS, à leurs négociations puis à la mise en place de la QoS, sont divisés en quatre catégories :

- **SCOPE**(ip_src, ip_dst, realm, rev_id) : Ces paramètres sont obligatoires et définissent la portée de la réservation. Le scope est composé essentiellement de l'adresse source (ip_src) et destination (ip_dst) et éventuellement de l'adresse « realm » en cas d'utilisation de la translation d'adresses (NAT). Ces adresses sont fixées lors de la requête initiale et ne peuvent être modifiées. Suite à une demande acceptée, il contient l'identifiant de réservation unique (rev_id) attribué par la fonction de CAC et qui sert d'identifiant pour les fonctions de modification, de suppression et d'information.
- **QSPEC**(bw, cos, tos, delay, jitter, loss) : Il définit la spécification de la QoS requise par l'application. Il contient au minimum la bande passante (bw) et la classe de service (cos). Si nécessaire, via le type de service (tos), il est possible de définir plus finement les paramètres de QoS (delay, jitter, loss). Il est possible de spécifier une bande passante minimale et maximale afin d'obtenir un débit entre ces deux valeurs plutôt qu'un refus, lorsqu'il n'y a plus assez de bande passante.

- **FSPEC**(ip_src, ip_dst, port_src, port_dst, label, mark, protocol) : Optionnels, ces paramètres permettent de spécifier l'enveloppe du flux « Flow Specification » auquel la QoS doit s'appliquer. Il est parfois nécessaire de définir précisément ces paramètres lorsque la fonction de CAC doit également configurer les équipements du réseau.
- **TSPEC**(start, stop, length, repeat) : Également optionnelle, la spécification du temps permet de préciser quand la QoS doit s'appliquer soit entre un début (start) et une fin (stop), soit sous forme de durée (length) et éventuellement peut être répétée (repeat). Elle permet surtout pour la fonction de CAC de révoquer toute réservation au bout d'un temps donné (stop ou length) afin de supprimer toute réservation fantôme qui pourrait survenir suite à une panne.

Le protocole doit pouvoir supporter aussi bien IPv4 que IPv6. Aussi, les adresses IP des champs SCOPE et FSPEC sont adaptées en fonction de la version du protocole IP. Dans le cas d'IPv6, le « flow label » servira avantageusement de paramètre d'identification des flux. Si la translation d'adresse (NAT) est supportée par le service, il est impératif que la désignation des adresses IP dans le champ scope soit unique. Cependant, pour des raisons de simplification, il n'est pas toujours possible de garantir l'unicité du plan d'adressage lors de l'utilisation de l'adressage privé¹¹.

4.1.2. Fonctions

Les fonctions de base associées au service Qu@ds sont les suivantes :

- **RESERVE**(scope, qspec) : Réserve des ressources dans le réseau auprès de la fonction de CAC. Elle comporte au minimum la portée (scope) et la qos (qspec). Il est possible de lui adjoindre la description du flux (fspec) et du moment (tspec) où la réservation doit s'appliquer. Si le temps n'est pas spécifié, une valeur de durée par défaut est attribuée par la fonction de CAC au bout de laquelle la réservation sera automatiquement supprimée. En retour, la fonction retourne l'identifiant de réservation et la spécification de QoS mise à jour.
- **DELETE**(scope) et **DELETE_ALL** : Supprime une réservation, ou toutes les réservations, sur la base de l'identifiant de réservation préalablement obtenu lors d'une admission d'appel réussie. La destruction de toutes les réservations est nécessaire en cas de problème ou de redémarrage après une panne afin de repartir sur un état stable.
- **MODIFY**(scope, qspec, fspec, tspec) : Modifie une réservation en cours sur la base de l'identifiant de réservation unique. Tout est modifiable excepté la portée (scope) de la réservation. En cas d'échec, la réservation précédente est maintenue. L'identifiant de réservation n'est pas modifié.
- **GET**(scope) et **GET_ALL** : Ces deux fonctions permettent de récupérer les informations liées à une réservation ou toutes les réservations. La dernière est utile pour la resynchronisation après une panne.

Toutes ces fonctions doivent être traitées de façon atomique, même en cas de demandes multiples, de façon à garder cohérentes les réservations stockées dans la base de données de la CAC. De plus le service étant transactionnel, une réservation multiple (bidirectionnelle ou point à multipoints), doit être traitée globalement. En cas d'échec d'une des demandes de réservations unitaires au sein de la requête multiple, toutes les réservations unitaires ayant abouti précédemment doivent être supprimées avant de rendre un avis négatif au demandeur.

11 Par exemple, tous les réseaux domestiques derrière une « box » xDSL, utilisent le même plan d'adressage

4.1.3. Propriétés de la couche de transport

Bien évidemment, la couche de transport doit être fiable et sans perte. TCP ou STCP étant de facto les protocoles à utiliser. Une couche d'abstraction supplémentaire peut être utilisée comme HTML/SOAP ou DIAMETER. Le protocole doit impérativement être transactionnel afin de garantir que toutes les demandes simultanées seront bien traitées en temps réel et que chaque transaction entraîne toutes les mises à jour de la fonction d'admission d'appel et de réservation. Il doit mettre en œuvre des stratégies particulièrement poussées de traitement des pointes de charge, des incidents possibles, des demandes simultanées d'écriture, tout en prenant toutes les mesures possibles pour préserver à tout moment la cohérence de la base de données. En cas de panne, Le protocole de transport doit pouvoir détecter une rupture de connexion avec le service et fournir des procédures prédéfinies de retour en arrière à un état connu et cohérent, avec annulation des dernières transactions et resynchronisation des états entre les logiques de service et le service [Qu@ds](#).

4.2. Limitations du Flow Spec

Contrairement à l'ATM ou MPLS qui permet d'identifier à coups sûrs les cellules ou les paquets appartenant à la même connexion, les paquets IP d'un même flux ou micro flux sont beaucoup plus délicats à identifier. Même le « Flow Spec » ainsi défini permet de spécifier tous les paramètres nécessaires à l'identification d'un flux ou micro flux, sa mise en œuvre est toutefois limitée et compliquée. En premier lieu, les numéros de ports sources et destinations ne sont pas toujours connus au moment de la requête de QoS. En particulier, le port source est systématiquement calculé dynamiquement par l'application cliente ce qui impose que la demande de QoS émane du client et non du serveur. Dans le cas de SIP et RTP, ces deux ports sont négociés lors de l'échange préalable des messages SDP entre le client et le serveur. Enfin, dans le cas des paquets fragmentés e.g. diffusion Vidéo en UDP, seul l'en-tête IP est reproduit dans les paquets fragmentés suivants. Les informations de ports sont ainsi perdus empêchant leur utilisation pour l'identification des paquets. La translation d'adresses dans les NAT et le passage au travers de tunnel IPsec ou VPN compliquent la reconnaissance à partir des adresses IP puisqu'elles changent au fur et à mesure que le paquet progresse dans le réseau. Enfin, il peut être nécessaire d'appliquer le même traitement de QoS à plusieurs flux ou micro flux identifiés par des caractéristiques différentes e.g. Flux voix et vidéo d'une application de visioconférence.

Certains problèmes peuvent être résolus. Ainsi, il est possible d'identifier les paquets fragmentés à l'aide du numéro de paquet en place et lieu des numéros de ports. La négociation de port et le passage de NAT et FireWall ont été résolus dans l'architecture IMS : la demande de QoS est effectuée par le proxy SIP (P-CSCF) en association avec le SPDF du RACS qui ont la connaissance de toutes les informations et la possibilité de configurer les adresses ou les numéros de ports. Mais la mise en œuvre de ces mécanismes devient très compliquée voire inefficace en terme de performance. Il est donc souhaitable de limiter les paramètres d'identifications des flux et effectuer cette action au plus proche de la source pour éviter toute ambiguïté. Enfin IPv6, avec le « flow label » dans l'en-tête et l'absence de translation d'adresse, permettra de simplifier grandement ce processus d'identification des flux.

4.3. Sécurité, authentification et comptage

Si la négociation de QoS entre les plates-formes des opérateurs ne pose pas de soucis, il devient impératif de se prémunir contre des attaques et déni de services lorsque l'interface est mise à disposition des applications. Le protocole de transport doit permettre l'utilisation de

TLS pour crypter les paramètres et sécuriser les échanges.

Cependant, en préalable, il faut pouvoir authentifier le demandeur soit au niveau du terminal lui-même soit au niveau de l'utilisateur. Toute forme d'authentification est possible. Si l'utilisateur ou le terminal sont déjà authentifiés e.g. connexion PPP en ADSL, il n'est pas nécessaire de le réauthentifier. L'accès au service de QoS à la demande peut utiliser les principes de « Single Sign On » ou tout autre plate-forme d'authentification.

Il est légitime de se poser la question de la facturation une fois que la demande de QoS a été honorée. De longs débats au sein de la communauté et une vision différente selon les opérateurs ne permet pas de trancher cette question. Cependant, les opérateurs vont de plus en plus vers des tarifications au forfait (abonnement mensuel plus forfait de communications) plutôt que à l'acte (facturation à la communication exclusivement). Entre politique tarifaire et complexité de mise en œuvre, la tendance est donc de ne pas comptabiliser l'utilisation de la QoS session par session au niveau des offres de détails, mais de facturer les offres de gros vers les opérateurs tiers ou les fournisseurs de services.

4.4. Applications

Ce travail de recherche a servi de base à de nombreuses contributions de la part de l'auteur en standardisation (Voir §I.4.5). Il se retrouve en partie dans l'interface Gq' de l'ETSI TISPAN figurant au dessus du RACS, de l'interface Rs de l'ITU-T figurant au dessus du RACF. Il a été implémenté dans le projet EuQoS sous le nom de EQ-SAP « EuQoS Service Access Point » comme étant le point d'entrée des requêtes de QoS au « Ressource Manager » (voir §III.5) et comme point d'entrée du module de CAC dans la passerelle domestique (voir §III.7). Des protocoles similaires ont également été standardisés comme la spécification des paramètres de QoS de NSIS [115] ou comme dans UPnP QoS [116].

5. Admission d'appel à l'inter-domaine

En complément de la gestion de la QoS dans son réseau, si un opérateur veut garantir de la qualité de service de bout en bout, il sera forcément confronté à la gestion des ressources sur ses liens d'interconnexions avec les autres opérateurs et à la nécessaire collaboration avec ces derniers. Aussi, aux contraintes techniques, viennent s'ajouter les contraintes des politiques des opérateurs, la sécurisation des échanges, la protection et la confidentialité des données échangées. C'est dans ce contexte et en continuité des travaux des projets TEQUILA, AQUILA, CADENUS et MESCAL, que le projet EuQoS [117] a initié ses travaux afin de fournir une solution pour la gestion de la QoS de bout en bout et pour des réseaux hétérogènes.

L'architecture de EuQoS [118] repose sur les mêmes fondements que CADENUS (Cf. §III.3) : séparation entre plan de service et plan de contrôle lui-même séparé entre niveaux indépendant et dépendant de la technologie de transport, et partitionnement des réseaux. Les contributions de ce mémoire, décrites ci-après, portent à la fois :

- sur le concept de chemin à QoS garantie,
- l'architecture du système EuQoS basée sur une répartition des fonctions d'admission d'appel,
- l'architecture du gestionnaire de ressources et sa machine d'état,
- sur la solution d'approvisionnement du système EuQoS à l'aide de tunnels établis par MPLS-DiffServ-TE.

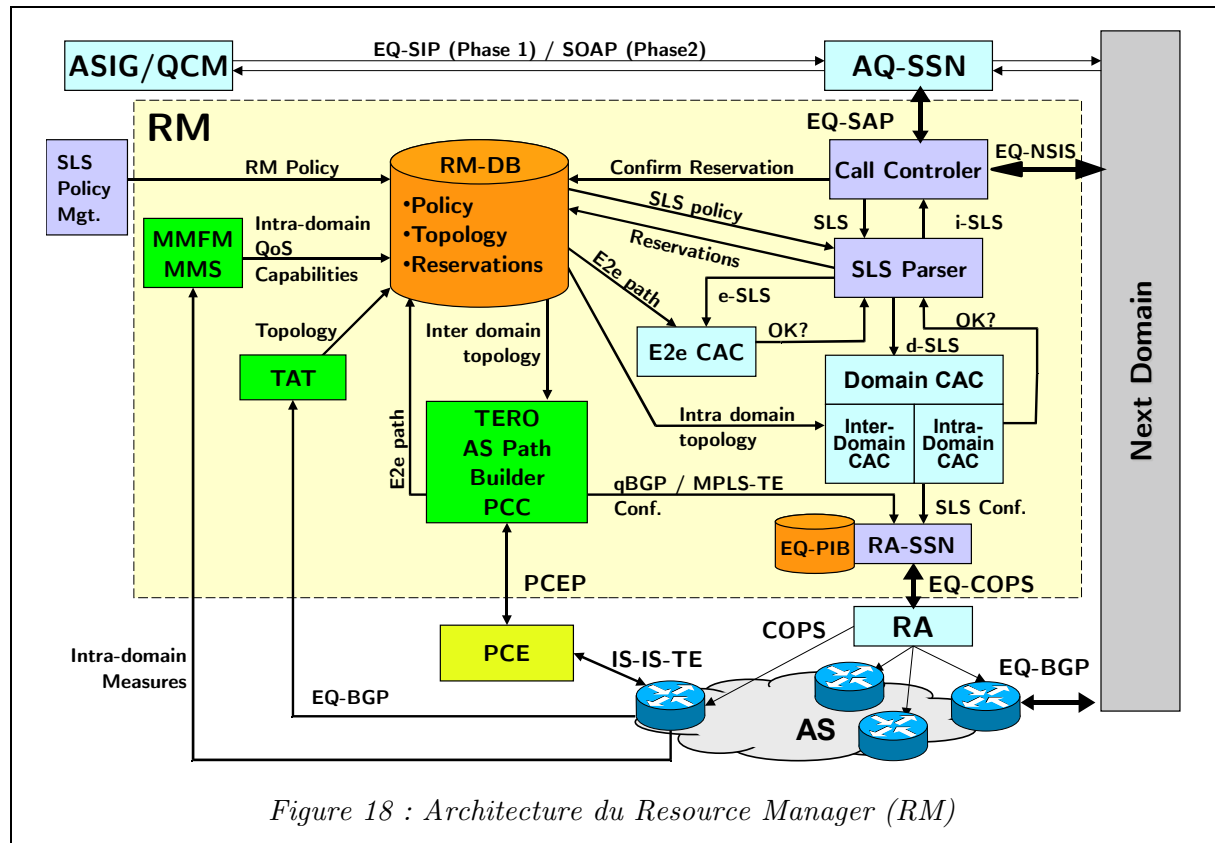
5.1. Architecture EuQoS

Le système EuQoS est divisé entre un plan de service et un plan de contrôle. Le plan de service regroupe toutes les fonctions nécessaires à la négociation de QoS entre les applications et le réseau. L'interface mise en œuvre se décompose entre le module « Quality Control Module - QCM » embarqué dans les applications et le module « Application Quality Service Signalling Negotiation - AQ-SSN » qui offre la possibilité de négocier la QoS. D'abord implémenté sous la forme de SDP sur SIP, la version définitive est basée sur SOAP offrant plus de flexibilité pour l'intégration au niveau des applications. Deux modules supplémentaires permettent d'authentifier les utilisateurs, « Service Authentification Autorization Accounting - SAAA » et de facturer les clients, « Charging - CHAR ».

Le plan de contrôle est lui-même divisé entre un plan indépendant de la technologie et un plan dépendant de la technologie. Ils sont respectivement pilotés par le « Resource Manager - RM » et le « Resource Allocator - RA » (voir figure 18). Ces deux composants sont en charge de négocier la QoS dans le domaine et entre opérateurs (RM) ainsi que s'assurer de la bonne configuration des équipements des réseaux sous-jacents (RA).

Enfin, au niveau du plan de transfert, EuQoS a étudié et défini un cadre commun à toutes les technologies de transports supportées (LAN, xDSL, WIFI, UMTS, MPLS, Satellite) afin d'homogénéiser les CoS et garantir une continuité de la QoS. Six classes de services de bout en bout ont été définies ; « Telephony », « Real-Time Interactive », « Signalling », « Multi-media Streaming », « High Throughput Data » et « Standard ». Les trois premières se voient offrir le même type de traitement des paquets (classe Real-Time) avec des objectifs stricts sur la perte, le délai et la gigue. Elles servent aux applications téléphoniques, de visio-conférences et de jeux interactifs, et pour le transport de la signalisation réseau. Les deux suivantes sont traitées en mode « Non Real-Time » avec les mêmes objectifs pour la perte et le délai mais en relâchant la contrainte sur la gigue. Elles servent aux applications de

streaming et de transfert de données. Finalement, la dernière classe correspond au mode « Best-Effort », donc, sans aucune garantie. Chaque classe est ensuite décilnée dans la technologie sous-jacente : DiffServ à l'inter-domaine et dans les réseaux cœurs, ATC et VC/VP séparés pour l'xDSL, extension WMM pour le WiFi, 802.1p et 802.1q pour le LAN, et une admission d'appel spécifique pour l'UMTS et le satellite.



5.2. Approvisionnement et utilisation des chemins à QoS

5.2.1. Définition des EQ-PATH

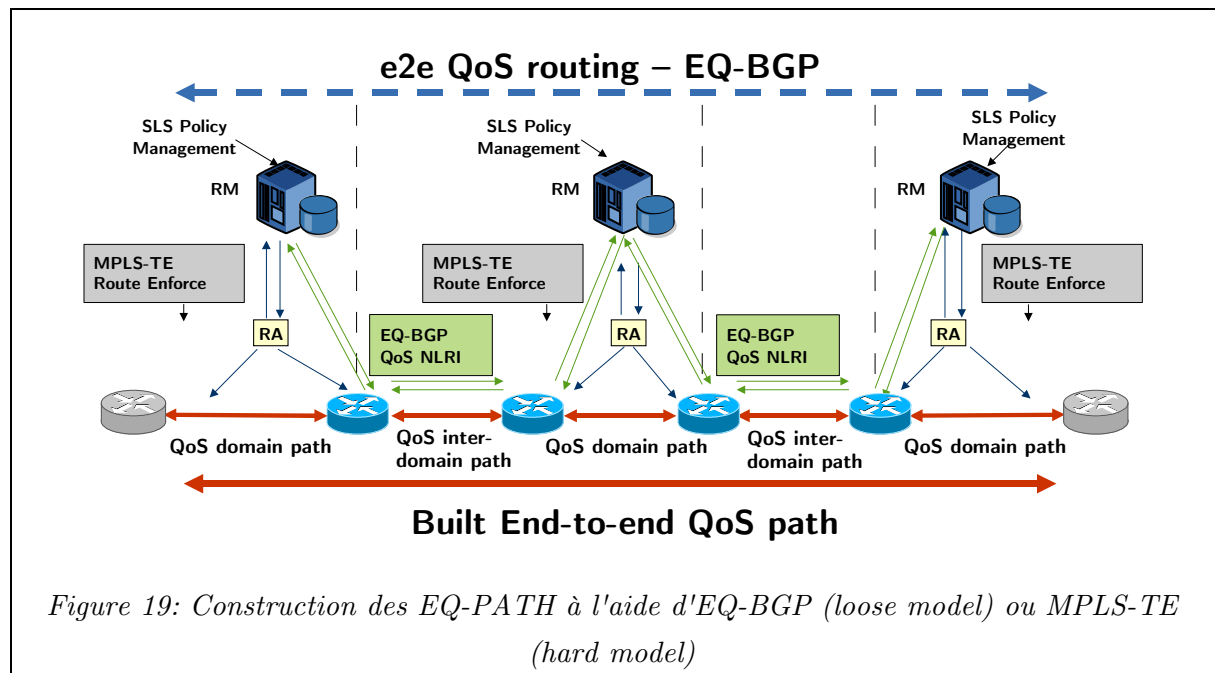
Le fondement de l'architecture [119] repose sur la construction de chemins de bout en bout à garantie de QoS appelés « **EQ-PATH** », durant le processus d'approvisionnement, dont l'usage est contrôlé par une fonction d'admission d'appel, durant le processus d'invocation, et surveillé durant le processus d'après-vente. Ces EQ-PATH sont construits par les RM localisés dans chaque partition de réseau à partir des informations de routage et des accords d'échanges entre opérateurs. Ces derniers sont formalisés sous la forme de p-SLS « peering SLS » et saisis par les opérateurs via l'interface de gestion du RM. Chaque EQ-PATH correspond à une classe de service supportée par le système EuQoS.

La création des EQ-PATH à l'inter-domaine passe par l'utilisation d'une version modifiée de BGP : EQ-BGP [120] qui, à la manière de qBGP, permet d'annoncer des informations de QoS à ses voisins. Ces métriques (perte, délai et gigue) sont annoncées pour chaque classe de service supportée. Ainsi, chaque routeur de bordures exécutant EQ-BGP apprend des routes qui peuvent être différentes en fonction de la classe de service.

5.2.2. Construction des EQ-Path

L'architecture du système EuQoS est bâtie autour du « Resource Manager » et du « Resource Allocator -RA ». Le RM se charge du traitement de la QoS indépendamment de la technologie de transport alors que le RA est implémenté en fonction de la technologie du réseau. Il existe 6 implémentations du RA, une pour chaque technologie supportée.

Une fois le p-SLS saisit dans le RM, il va être traité par le module d'ingénierie de trafic « Traffic Engineering and Route Optimization - TERO » qui se charge de configurer les routeurs de bordures avec les métriques de QoS qui seront annoncées par EQ-BGP. Une fois que EQ-BGP a convergé, dans un temps similaire à BGP, les tables de routages par CoS sont disponibles et sont injectées dans la base de données RM_DB du RM par le module de gestion « Monitoring Measurement and Fault Management » via le module d'acquisition de topologie « Topology Acquisition Tool - TAT ». Le RM dispose ainsi dans sa base de données, l'ensemble des EQ-PATH pour chaque classe de service et pour chaque destination. Dans ce mode de fonctionnement (voir figure 19), appelé « loose model », seules les relations avec les AS voisins sont réellement connues. C'est le protocole de routage BGP classique qui fournit les vecteurs de distances i.e. les chemins d'AS, qu'il faut suivre pour une destination donnée. L'EQ-PATH est donc construit et connu des RM de proche en proche comme dans l'Internet. Un deuxième mode de construction, appelé « hard model », configure des tunnels MPLS-TE pour construire les EQ-PATH. Les tunnels sont construits de proche en proche et permettent de contraindre le routage tout en réservant des ressources.

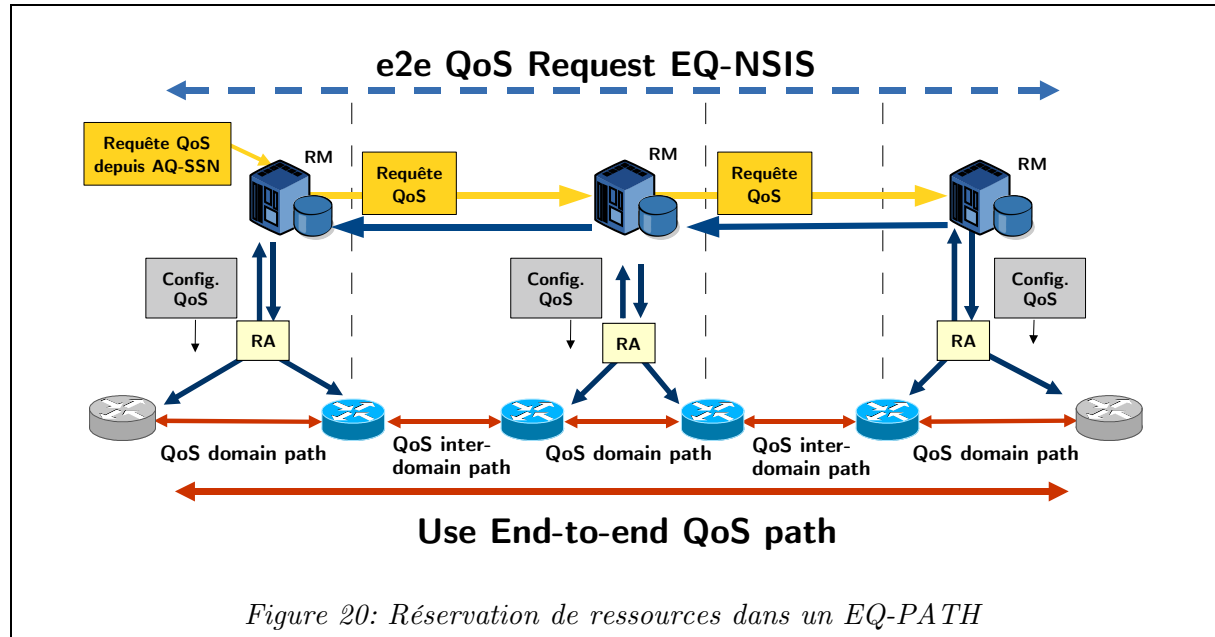


5.2.3. Utilisation des EQ-PATH

Une fois les EQ-PATH approvisionnés dans les différents réseaux, un utilisateur peut alors solliciter le système EuQoS pour obtenir une connexion avec QoS garantie. Deux modes pour cela selon que l'utilisateur utilise une application compatible avec le système EuQoS ou non. Dans le deuxième cas, il faut pouvoir disposer d'un proxy qui rend l'application compatible ou utiliser le service de base via le serveur Web. Ensuite, la requête de QoS est construite à partir des données fournies par l'application (ou le proxy, ou le serveur Web)

puis envoyée au RM via le protocole EQ-NSIS qui gère le réseau où le terminal est raccordé (voir figure 20).

L'interface d'accès au RM implémente un service de QoS à la demande, appelé EQ-SAP conforme à celui défini au chapitre §III.4. EQ-NSIS est une version étendue du protocole NSIS ayant la possibilité de travailler en mode « Off-Path ». Les paramètres de QoS sont conformes au qNSLP définie à l'IETF. Dans la première version du système EuQoS, les échanges étaient réalisés en SIP qui transportait en SDP spécifique contenant la requête de QoS. Le même protocole EQ-SIP est utilisé entre le RM durant la négociation de QoS.



Comme dans le médiateur de ressources de CADENUS, le contrôleur d'appel (module call-controller) du RM va extraire le SLS de la requête de QoS et le découper en entités correspondantes à chaque traitement. Il effectue un premier niveau d'admission d'appel de bout en bout (module e2e CAC). Il consiste dans un premier temps à déterminer si il existe un EQ-PATH à même de transporter ces données dans la classe de service et la destination spécifiées dans la requête. Ensuite le module d'admission d'appel à l'inter-domaine réserve le débit nécessaire pour cette connexion sur le lien inter-domaine, i.e. sur le routeur EQ-BGP de bordure. Une fois acceptée, la demande est traitée par le module d'admission d'appel intra domaine qui vérifie la disponibilité des ressources entre le terminal et le nœud de bordure. A ce stade, les différents niveaux d'admission d'appel ont permis de réserver des ressources dans le premier réseau et de sélectionner l'EQ-PATH à suivre. La demande est alors envoyée au RM suivant sur le chemin d'AS que suit l'EQ-PATH. La base de données qui contient les tables de routages contient également l'adresse IP des RM des AS voisins de sorte à ce que le RM sache à qui transmettre les requêtes de QoS en fonction des EQ-PATH. Chaque RM procède au même traitement que le premier excepté la sélection de l'EQ-PATH. Le dernier RM renvoie vers le premier RM la confirmation de la réservation finale. Sur réception de cette confirmation, chaque RM contacte alors son ou ses RA afin de configurer la QoS dans le réseau sous-jacent. En cas d'échec, le RM qui ne peut garantir la QoS, i.e. dont l'une des admissions d'appels a refusé la réservation, renvoie un message de libération au RM précédent et ne transmet pas la requête au RM suivant. L'admission d'appel est de type répartie et réalisée de proche en proche suivant le modèle cascade de CADENUS. Un RM

peut également contacter son RA tout en transmettant la requête au suivant afin de gagner du temps sur la mise en place de la QoS sans attendre la réponse finale. En cas de réponse négative, la configuration devra être annulée au même titre que la réservation initiale.

5.3. Construction des EQ-Path selon le « hard model »

Si cette architecture donne de bons résultats tant en simulation qu'avec les mesures effectuées sur les 12 plates-formes de démonstrations du projet EuQoS, il n'en demeure pas moins que le modèle « loose » ne permet pas de garantir une QoS absolue sur les liens inter domaines. A la fois, parce que le routage EQ-BGP (comme BGP) fluctue avec les nouvelles annonces de routes et parce que la réservation de débit n'est garantie que par le RM. Il n'y a pas d'isolation des EQ-PATH au niveau du routeur EQ-BGP, mais juste une isolation des classes de services via la mise en place de DiffServ. Afin de palier ces problèmes, l'architecture EuQoS a été améliorée en permettant la construction des EQ-PATH à l'aide de tunnels MPLS-DiffServ-TE [121]. Ce modèle d'architecture est appelé « hard » par opposition au précédent car il permet de contraindre le routage.

5.3.1. Définition des EQ-LINK

Les EQ-PATH désignent un chemin de bout en bout porteur d'une classe de service donnée, et sont construits de proche en proche par les routeurs de bordures de domaines « AS Border Router -ASBR ». Comme tous les réseaux n'implémentent pas ni ne supportent¹² MPLS, les tunnels qui sont créés dans le cadre du système EuQoS sont appelés des EQ-LINK. Par définition, un EQ-PATH est constitué de 0 à n EQ-LINK de sorte à rester compatible avec le modèle « loose » et les réseaux ne disposant pas de MPLS. Ainsi, le système EuQoS peut s'accommoder de réseaux gérés en mode « loose » ou « hard ». Un EQ-LINK garantit nativement la QoS du fait qu'il est construit comme un tunnel MPLS-DiffServ-TE. Ce dernier est associé à une file d'attente correspondant à la classe de service qu'il supporte, garantissant des objectifs en terme de délai, gigue et perte. De plus, durant son établissement, le protocole RSVP-TE réserve dans chaque routeur MPLS la bande passante requise pour ce tunnel. Trois types de tunnels inter-domaine ont été définis [122] par l'IETF [123] :

- Les tunnels contigus de bout en bout qui ont pour principal inconvénient d'obliger les opérateurs à exposer leur interface de bordures. Ceci présente des risques sur la confidentialité et la sécurité. La protection du tunnel impose les mêmes contraintes au tunnel de backup,
- Les tunnels aboutés par morceaux dit « stitched » qui permettent de s'affranchir partiellement de ces contraintes. Chaque opérateur est responsable de son domaine et seul le numéro de label entre opérateurs a besoin d'être connu. Les tunnels sont indépendants les uns des autres tout en formant une continuité d'acheminement. Ils sont cependant plus difficiles à protéger,
- Les tunnels hiérarchiques ou « nested » [124] qui résolvent les problèmes précédents. Les labels sont établis indépendamment par chaque opérateur. Les tunnels en provenance d'un AS sont ensuite agrégés (empilés au sens des labels) dans un tunnel de niveau hiérarchique supérieur à destination de l'AS suivant. La protection est plus simple puisque gérée localement par chaque opérateur.

5.3.1. Construction des EQ-LINK

L'établissement des EQ-Link nécessite le calcul du meilleur chemin d'AS en fonction de

¹² Cas des réseaux radio

la classe de service supportée puis le calcul du placement du tunnel dans chacun des réseaux. Le calcul du placement des tunnels est généralement effectué par le routeur de tête du tunnel. Si celui-ci à la connaissance de son réseau, il n'a aucune visibilité sur les réseaux suivants et ne peut donc pas calculer efficacement un placement de tunnel inter-domaine. Le même problème apparaît lors de l'utilisation du multi aire avec OSPF. L'IETF a commencé à standardiser le « Path Computation Element - PCE » pour calculer efficacement le placement des tunnels MPLS.

Le principe du PCE est de calculer le placement optimal des tunnels sur un réseau en fonction de la disponibilité des ressources physiques annoncées par le protocole de routage à ingénierie de trafic du réseau : IS-IS-TE [125] ou OSPF-TE [126]. En effet, le PCE participe au protocole de routage en tant que source passive. Ceci permet à la fois d'annoncer où sont localisés les PCE dans le réseau et au PCE de stocker les informations d'ingénierie de trafic, essentiellement la bande passante disponible sur les différents routeurs du réseau, dans sa base de données « Traffic Engineering Database - TED ». La TED contient ainsi toutes les informations, à jour, relatives à la topologie du réseau nécessaire au placement des tunnels.

Une fois le calcul terminé, le routeur de tête du tunnel initialise la signalisation RSVP-TE vers le routeur de fin du tunnel avec la route calculée par le serveur PCE (option « Explicit Route Object : ERO » de RSVP-TE). Il est ensuite automatiquement annoncé par le protocole de routage qui simultanément met à jour la bande passante restante dans le réseau. Ainsi, lors du prochain calcul de chemin, le PCE est automatiquement averti de la nouvelle disponibilité de ressources. Ceci permet de disposer d'un serveur PCE sans état (state-less) bien qu'il soit possible d'implémenter une version avec maintien des états. La seule contrainte de ce modèle est de ne pas solliciter de nouveau le PCE avant que le tunnel ne soit installé et annoncé dans le réseau. Cependant, le PCE peut accepter une requête portant sur plusieurs placements de tunnels, ce qui lui permet d'optimiser le placement de tous les tunnels simultanément plutôt que un par un.

Dans le cas de tunnels traversant plusieurs AS, le PCE déclenche l'algorithme de calcul récursif de chemin « Backward Recursive Path Computation -BRPC » [127] et contacte le PCE qui gère l'AS suivant. De proche en proche, le PCE de l'AS de destination reçoit la requête et commence à calculer le placement du tunnel entre le routeur de fin et les routeurs de frontières avec le domaine précédent. Il peut exister plusieurs routeurs de bordures ASBR entre deux domaines¹³. Le dernier PCE va ainsi renvoyer au PCE précédent autant de routes, appelée « Explicit Route Object - ERO » qu'il existe d'ASBR. À sa réception, le PCE précédent va également commencer son calcul de placement de tunnel mais cette fois-ci entre les ASBR de sorties (fournis comme points d'entrée du domaine suivant dans l'ERO) et les ASBR entre lui et le domaine précédent. Une telle procédure de calcul récursif depuis la destination, devient rapidement exponentielle vis à vis du nombre de solutions possibles. Afin de réduire cette combinatoire, l'algorithme du BRPC stipule que chaque PCE ne retient qu'un seul chemin, le plus performant, par ASBR de sortie. Ainsi, le nombre de solutions reste borné par le nombre d'ASBR entre deux domaines. De proche en proche le premier PCE reçoit l'ERO contenant les solutions possibles pour atteindre la destination. Il calcule à son tour le placement du tunnel entre la source et ses ASBR vers le domaine suivant. Puis, il choisit le meilleur des chemins de sorte à ce qu'il n'en reste plus qu'un. Une fois l'ERO calculé de bout en bout, la mise en place du tunnel s'effectue à l'aide de la signalisation RSVP-TE à partir du routeur de tête. Afin d'accélérer le calcul, chaque PCE peut anticiper le calcul du placement de ses tunnels avant d'avoir reçu l'ERO du PCE suivant.

¹³ Pour des besoin de partage de charge ou de sécurisation

Cependant, l'IETF ne recommande rien sur le calcul du chemin d'AS. L'architecture EuQoS associe la mécanique PCE avec le module TERO du RM (voir figure 21). Lorsqu'un opérateur saisit un nouveau p-SLS, les modules TERO de chaque RM calculent le meilleur chemin d'AS possible, à partir des données issues du protocole EQ-BGP. Le contrat est tout d'abord saisi par l'opérateur de destination pour sécuriser l'établissement afin d'éviter des demandes de calculs non sollicitées. Le calcul (1) s'effectue selon le principe du « source routing » [128] en tenant compte des objectifs de QoS de la classe de service, de la capacités des différents domaines et des contraintes administratives liées aux p-SLS. Ensuite, le module TERO sollicite (2) le PCE de son domaine pour le calcul de placement du tunnel en lui fournissant le chemin d'AS calculé précédemment. Les PCE coopèrent selon l'algorithme du BRPC (3) pour calculer le placement du tunnel MPLS-DiffServ-TE inter domaine. En retour, le PCE lui procure l'ERO du tunnel (4). Il ne reste plus qu'au module TERO de solliciter le routeur de tête, via le RA MPLS (5), pour qu'il déclenche la construction du tunnel à l'aide du protocole RSVP-TE muni de l'ERO (6). Pour finir, le protocole de routage à ingénierie de trafic IS-IS-TE met à jour la TED (7) des PCE pour tenir compte de ce nouveau tunnel. Une fois les EQ-LINK établies, les EQ-PATH sont construits comme une succession de 1 à n EQ-LINK et sont stockés dans la base de données du RM.

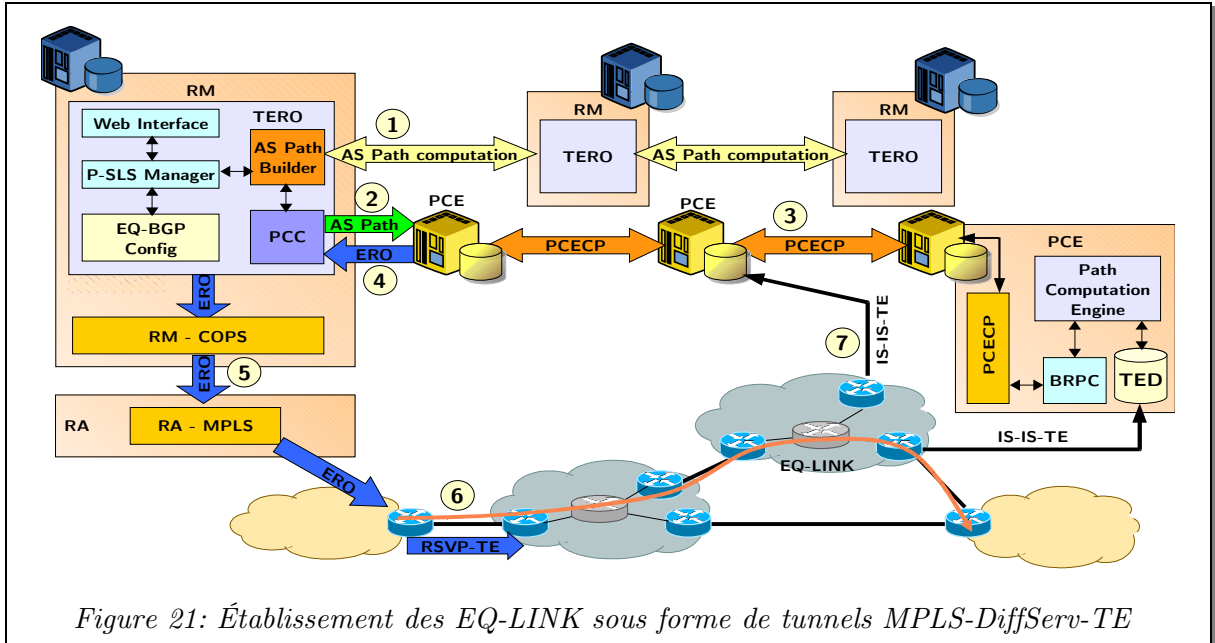


Figure 21: Établissement des EQ-LINK sous forme de tunnels MPLS-DiffServ-TE

Bien que la distance moyenne dans l'Internet soit comprise entre 3 et 5 AS [129], l'établissement d'un réseau maillé de tunnel entre les opérateurs ne peut s'appliquer si le nombre d'AS augmente significativement. Il est donc préférable d'utiliser des tunnels hiérarchiques pour construire les EQ-LINK afin de mieux supporter le passage à l'échelle. Le nombre d'EQ-LINK gérés par chaque RM est alors borné par le nombre de nœuds de bordures. Un maillage complet d'EQ-LINK entre tous les nœuds de bordures pour toutes les classes de services est donc tout à fait gérable.

5.3.2. Utilisation et protection des EQ-LINK

Lorsque le premier RM reçoit une demande de QoS, il cherche dans sa base de données l'EQ-PATH correspondant à la CoS demandée comme énoncé précédemment. Les différentes phases d'admissions d'appels se déroulent de façon similaire, puis le contrôleur d'appel contacte le prochain RM impliqué dans la gestion de l'EQ-PATH et fourni par la base de

données. La différence vient que l'adresse du prochain RM ne correspond pas au RM du domaine suivant mais au RM du domaine destinataire de l'EQ-LINK, i.e. le RM gérant la fin du tunnel MPLS. Ainsi, l'EQ-LINK peut être vu comme un raccourci entre deux RM distant de plusieurs AS.

Dans le cas d'EQ-PATH formé exclusivement par des tunnels hiérarchiques, seuls les deux RM d'extrémités sont impliqués dans la réservation de débit pour l'application requérante. En effet, il suffit de contrôler uniquement le volume de trafic entrant dans le tunnel. Il n'est pas nécessaire de vérifier dans les domaines de transit si la bande passante est disponible puisqu'elle a été réservée durant l'établissement du tunnel via RSVP-TE. Ce principe permet d'améliorer les performances globales du système EuQoS en réduisant le nombre de RM impliqués dans le traitement de la signalisation. La configuration effectuée par les RA est également limitée. Il suffit de configurer le nœud d'accès pour qu'il marque les flux de l'application avec le code DiffServ adéquat, avant que les paquets n'entrent dans le tunnel.

5.4. Bilan et Conclusion

Le projet EuQoS, au travers des 12 réseaux de tests interconnectés via GEANT, a prouvé que l'architecture permet de garantir de la QoS de bout en bout tout en étant réalisable et applicable à des réseaux de grande taille et à de nombreuses technologies [130]. La séparation en processus, en couche et en partition rend la solution flexible et extensible. Le support d'une nouvelle technologie de transport, e.g. GMPLS pour les réseaux optiques, se réalise simplement en développant seulement le RA adapté. Il est également possible de diviser un réseau de trop grande taille en plusieurs partitions e.g. division d'un réseau d'accès en PoP fédérés autour d'un réseau cœur, d'assigner un RM par partition puis d'utiliser l'interface inter-domaine des RM à l'échelle des partitions. L'API de développement de l'interface EQ-SAP a permis de rendre compatibles des applications aussi variées que la visioconférence, la téléphonie, les jeux ou la vidéo à la demande. Le système a également été testé auprès de médecins avec une application de télé-médecine.

Cependant, le déploiement de la solution EuQoS passera forcément par une phase de standardisation et de mise à disposition dans les équipements du protocole EQ-BGP. Malheureusement, le protocole qBGP ayant été confié à l'IRTF, la standardisation des métriques de QoS pour BGP est repoussée de fait, à plus tard. Certes, le consortium EuQoS a validé l'utilisation d'EQ-BGP entre « route reflector » au-dessus d'un réseau constitué de routeurs standards exécutant BGPv4, mais la solution perd en flexibilité. Aussi, pour un déploiement et une adoption rapide, le modèle « hard » est préféré car il n'impose pas de disposer d'EQ-BGP au niveau des routeurs de bordures. Les modules TERO doivent juste disposer des informations de QoS émanant des différents AS. Les EQ-PATH sont matérialisés ensuite par la construction des tunnels MPLS-DiffServ-TE plutôt que par un routage à QoS garantie. Les mesures effectuées sur le démonstrateur EuQoS ont donné d'excellents résultats puisque le temps maximale d'établissement d'un EQ-LINK reste inférieur à 5 secondes. L'architecture EuQoS est ainsi proche de la solution finale exposée au §III.8, à laquelle il ne manque que l'interconnexion avec le réseau domestique.

6. Acquisition de topologie

Comme exposé dans les chapitres §II.7 et §III.3, la fonction d'admission d'appel et de réservation de ressources nécessite impérativement d'être synchronisée avec le plan de transfert afin de disposer de la topologie réelle du réseau. Deux méthodes sont possibles :

- Soit le plan de contrôle contraint le routage comme proposé dans le chapitre précédent à l'aide de MPLS-DiffServ-TE. La topologie est alors connue de la fonction d'admission d'appel, puisqu'elle imposée par la plate-forme de contrôle,
- Soit le plan de transfert est autonome et il est alors nécessaire d'acquérir la topologie.

S'il existe des techniques, comme l'écoute des protocoles de routages IP ou à base de sondage [131], des protocoles comme le « Cisco Discovery Protocol - CDP » [132] et des produits d'acquisition de topologie basé sur SNMP [133], ils sont soit incomplets soit inadaptés. Ainsi, la couche 2 des réseaux (ATM, GigaBit, VLAN, MPLS, ...) n'est pas détectée alors que la gestion des ressources passe par la vérification des points de contention des réseaux aussi bien aux niveau 2 que 3, comme démontré dans le cadre du projet EuQoS. De même, la mise à jour en temps réel des modifications dans le réseau de transport n'a pas de solution lorsqu'il n'y a pas présence d'un protocole de routage IP.

Le travail de recherche sur l'acquisition de topologie a été mené dans le cadre de la thèse de 3^{ième} cycle de Walid Htira. Outre l'encadrement et l'orientation des travaux, la contribution de l'auteur de ce mémoire concerne le concept de l'automatisation de l'acquisition de la topologie y compris la synchronisation des changements et le calcul des contraintes pour l'admission d'appel sur le schéma en étoile à l'aide d'un nucléus réel.

6.1. STAMP

L'idée majeure est de s'inspirer des protocoles comme LLDP [134] de l'IEEE qui laisse les équipements s'auto annoncer plutôt que d'essayer de découvrir la topologie de proche en proche par sondage. Le protocole STAMP [135] « Scalable Topology Announcement and Management Protocol » a été conçu dans le but de remonter les informations de topologie d'un réseau par les nœuds le composant. Il repose sur trois concepts fondamentaux : une découverte des adjacences liées à la technologie par chaque équipement, un modèle de données basé sur un schéma d'encapsulation générique des différents protocoles de transport et une agrégation des informations collectées pour reconstituer la topologie. De plus, STAMP permet de récupérer les informations de QoS (débit des interfaces, classes de service, débit par classe, ...) nécessaires à la fonction de CAC.

Le protocole opère entre les équipements du réseau, les agents, et un collecteur de topologie, le manager, au sein d'un même groupe. L'appartenance à un groupe s'effectue par configuration suite à une découpe administrative, ou à partir de la découpe existante issue d'un protocole de routage, par exemple une aire OSPF, une région IS-IS, un peer-group ATM ou un domaine IP (Autonomous System). La phase de découverte de l'algorithme se déroule alors en deux temps :

- Chaque équipement découvre ses adjacences en envoyant des messages HELLO qui contiennent ses propres informations de topologie dont son identifiant unique (adresse MAC la plus faible) et les informations liées à l'interface utilisée pour envoyer le message. Il obtiendra de son voisinage des messages HELLO comportant le même type d'information. Ce premier niveau d'échange permet à la fois de découvrir ses adjacences mais également les informations de topologie de ses voisins en terme de connectivité.

- À partir de l'ensemble des adjacences plus des informations propres à l'équipement (table de routage, de commutation, listes des interfaces, adressage ...) le nœud reconstitue son morceau du puzzle. Une fois connecté puis identifié auprès du collecteur de topologie, l'équipement envoie ses données de topologie. La plate-forme responsable de la collecte des informations utilise alors l'identifiant unique de chaque équipement pour reconstituer la topologie complète à partir de l'ensemble des morceaux reçus.

Le protocole est également découpé en 2 couches :

- Une première couche en charge du niveau 2, liaison, qui permet la découverte des adjacences pour une technologie donnée. Deux technologies sont supportées à ce jour : l'ATM et Ethernet. La modularité de l'approche permet d'envisager facilement l'ajout de technologie supplémentaire comme le POS,
- Une deuxième couche gérant le niveau 3 qui collecte à la fois les informations de routages et communique les informations de topologie au collecteur. Basée sur TCP/IP, cette partie est commune aux agents tournant sur les équipements et au collecteur.

La remontée des informations de topologie est effectuée selon un codage « Type Length Value - TLV » bien connu en standardisation. Outre les TLV d'identification et de description de la technologie, chaque TLV va décrire un type de données correspondant au schéma d'encapsulation des protocoles de transport. L'ensemble des données est ensuite stocké dans différentes bases de données : une pour la gestion des adjacences, une pour la reconstitution complète de la topologie et une pour la connectivité qui reproduit le routage et/ou la commutation.

Par la suite, les équipements envoient à intervalle régulier des messages HELLO pour vérifier s'il n'y a pas de changement avec leur voisinage. Si un message HELLO n'est pas reçu ou si une alarme est détectée, l'équipement met à jour sa topologie et envoie au collecteur la modification. La plate-forme en charge de la reconstitution de la topologie sera avertie immédiatement si un problème est détecté.

Cette solution est supérieure à l'état de l'art dans le domaine dans le sens qu'elle répond aux critères de complétude, de précision, de réactivité, d'efficacité et de rapidité nécessaire à son utilisation pour alimenter une fonction d'admission d'appel centralisée.

6.2. Agrégation de topologie

Une fois la topologie apprise par la fonction d'admission d'appel, il reste deux problèmes de performance à résoudre. Tout d'abord, sur des réseaux de grande taille, la recherche du chemin suivi par le flux requérant de la QoS peut s'avérer longue et fastidieuse. Le parcours dans la base de données de topologie est du même ordre de grandeur que le parcours d'un graphe à la recherche du plus court chemin.

Sur une base de 10 requêtes de services par heure maximum, une durée de service moyenne de 10 minutes, un taux moyen d'utilisation de 10% et pour 10 Millions d'utilisateurs, le nombre de requêtes de QoS par seconde, c'est à dire le nombre d'appels par seconde, s'établit à 2777 appel/s. Ce qui donne un temps de service moyen de 360 μ s par appel tout compris. Le temps de parcours du graphe doit nécessairement être court sachant qu'il faut déjà repérer les points d'entrée (adresse IP source) et de sortie (adresse IP destination) dans le graphe.

Mais le plus pénalisant pour les performances intervient en cas de modification de la topologie (panne, routage, ...) qui impose à la plate-forme de CAC de recalculer toutes les

réservations déjà effectuées afin de mettre à jour sa base de données en fonction de la nouvelle topologie. Sur la même base de dimensionnement, la fonction de CAC doit gérer environ 500 000 réservations simultanées. Au maximum, si toutes les réservations sont impactées par la modification de topologie, le temps de re-calcule est égal à 3'45". Pendant cette période, la plate-forme ne peut admettre de nouvelles requêtes car chaque opération d'admission doit être traitée de façon atomique, y compris les mises à jour. Si la topologie varie trop, ce qui peut être le cas lorsqu'une mise à jour des tables BGP de l'Internet sont annoncées dans le réseau¹⁴, la fonction d'admission passe tout son temps à mettre à jour ses réservations et ne peut plus en admettre de nouvelle.

Pour finir, les opérateurs utilisent souvent dans leur réseau le partage de charge qui consiste à établir plusieurs liens entre deux routeurs afin de répartir le trafic. Le même principe est également appliqué sur plusieurs routeurs pour sécuriser le réseau. Une telle topologie maillée est très difficile à appréhender pour la fonction de CAC du fait qu'elle ne sait pas précisément par où passe le trafic à un instant t . Le partage de charge étant dynamique, ce sont les routeurs qui décident sur quel lien envoyer un flux.

Afin de pallier ces inconvénients, le mémoire se propose d'agréger les informations de topologie afin de réduire la volumétrie de la topologie et simplifier le processus d'admission d'appel. Plusieurs schémas d'agrégation ont déjà été proposés [136] pour déterminer le meilleur chemin dans un réseau en fonction de la QoS. L'idée principale est d'appliquer les mêmes principes pour l'admission d'appel en supprimant l'étape de parcours de la topologie selon le graphe de routage [137]. Seul les points d'entrée et de sortie du réseau, i.e. les nœuds de bordures, sont conservés et la réservation est effectuée entre ces deux points. Cependant, la déformation introduite par le processus d'agrégation sur le partage des liens internes du réseau, ne permet pas d'obtenir un taux de confiance suffisant pour la fonction de CAC. Des contraintes supplémentaires à vérifier ont été introduites dans l'algorithme de CAC sur chacun des deux modèles de schémas étudiés : l'étoile (star) [138] et le maillage complet (full-mesh) [139].

L'algorithme ainsi obtenu sur les schémas en étoile et maillé, donne d'excellents résultats en terme de performance, par rapport au schéma détaillé, puisque inférieure de l'ordre de 3 à 10 fois suivant les topologies [140]. Le temps de décision de l'algorithme de CAC est compris entre 10 et 250 μ s avec le modèle « star » selon la topologie et le nombre de paramètres de QoS pris en compte en comparaison de 100 à 6200 μ s sur le schéma détaillé tout en conservant une marge d'erreur inférieure à 2%. Mais le gain le plus significatif intervient lors d'une modification de topologie. En effet, sauf disparition d'un nœud de bordure, toute modification de routage se traduit, après calcul de l'agrégation, à une simple mise à jour des règles à vérifier dans l'algorithme de CAC. La mise à jour des réservations est alors immédiate puisqu'il suffit de vérifier la capacité des seuls liens agrégés entre nœuds de bordures qui ont été modifiés à l'issue de ce changement de topologie. A nouveau le gain par rapport au modèle classique est de l'ordre de 3 à 10 suivant les topologies et le schémas d'agrégation. De plus, les modifications mineures de routage peuvent passer inaperçues à l'issue de l'étape d'agrégation, et laisser le schéma agrégé inchangé.

6.3. Conclusion

Si les solutions proposées [141] permettent de résoudre les leviers technologiques que sont l'acquisition de topologie et les performances des plates-formes de CAC, il n'en demeure pas moins qu'il reste du chemin à parcourir avant de pouvoir déployer la solution dans un réseau opérationnel. En premier lieu, l'agent STAMP doit nécessairement être implémenté

14 De l'ordre de 50000 à 300000 update BGP par jour (mesure 2002)

par les fabricants d'équipements. Une piste d'accélération pour l'adoption du principe, serait d'utiliser le récent protocole standardisé par le groupe de travail IPFIX de l'IETF qui permet de remonter vers un manager des informations de mesures (à la NetFlow de Cisco), mais qui a été conçu de façon générique. Le protocole émet d'abord le schéma de représentation des données, puis les données elles-mêmes. Il est tout à fait envisageable de réutiliser IPFIX pour transporter les informations de topologies entre les agents STAMP et le manager.

Finalement, la généralisation de l'usage de tunnels MPLS-DiffServ-TE¹⁵ dans les réseaux permet de s'affranchir à la fois de la remontée d'informations et stabilise la topologie utilisée par la fonction de CAC. En effet, le tunnel simplifie la topologie car il matérialise le lien entre nœuds de bordure et n'est pas sensible aux fréquentes annonces de route BGP. De plus, si les tunnels de secours sont déployés en même temps, une panne sur le réseau est complètement transparente pour la fonction de CAC. Le « Fast Re-Route » permet de basculer sur le tunnel de secours en moins de 50 ms. Ce tunnel de secours ayant les mêmes caractéristiques de QoS que le tunnel primaire, la mise à jour des réservations empruntant ce tunnel n'est pas nécessaire. En revanche, le modèle agrégé peut servir à planifier le déploiement des tunnels et aider le calcul de placement des tunnels en cherchant à minimiser les équations gérant le partage des liens internes.

15 Cette architecture est détaillée dans la dernière partie de ce mémoire

7. Admission d'appel dans le réseau domestique

Les opérateurs de télécommunication basent aujourd'hui toute leur stratégie sur les nouvelles offres de services à valeur ajoutée. Outre l'accès à Internet, ils proposent maintenant des bouquets de services incluant la TV, la Vidéo à la Demande et la Voix sur IP appelés « Triple Play ». Pour les opérateurs à la fois fixes et mobiles, c'est un bouquet « Quadruple Play » qui est offert en ajoutant la convergence entre fixe et mobile permettant de profiter de son accès fixe pour passer des appels depuis son portable. L'ensemble de ces services est offert majoritairement sur un accès xDSL (ADSL, ADSL2+, VDSL, ...) et sur fibre optique quand elle est disponible. Comme l'ensemble de ces services est vendu au forfait, il est impératif pour un opérateur d'offrir une excellente qualité de service afin de se démarquer de la concurrence. S'il est possible de surdimensionner ou de gérer avec DiffServ les réseaux de cœur et de collecte, les mêmes mécanismes ne peuvent pas ou difficilement être appliqués aux réseaux d'accès. Il reste ainsi deux points majeurs de congestion dans le réseau : la ligne de cuivre et le lien de raccordement au réseau de collecte, c'est à dire de part et d'autre du DSLAM qui matérialise la frontière entre les 3 réseaux, domestique, accès et collecte. Les contributions de l'auteur dans ce domaine sont :

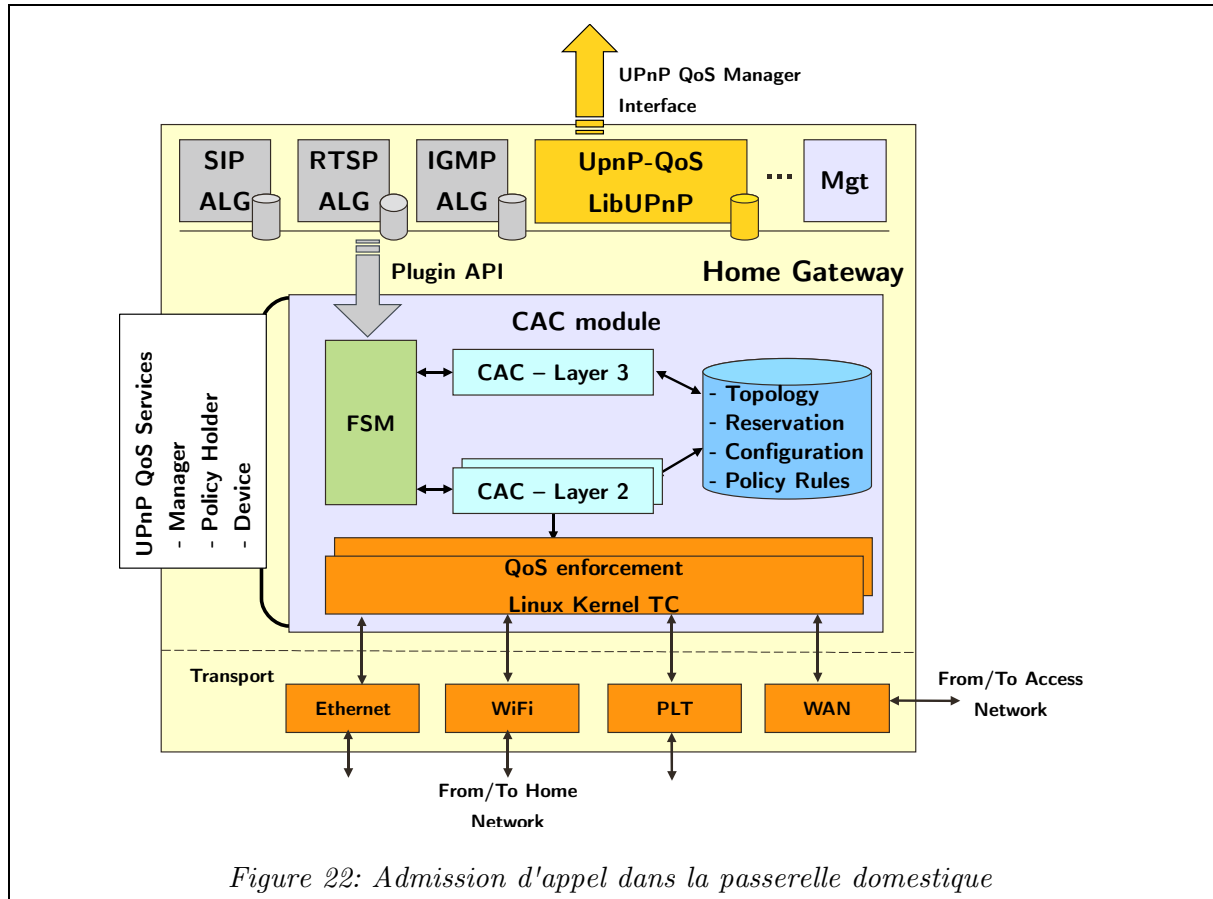
- la gestion de la QoS dans le réseau domestique à l'aide d'une fonction d'admission d'appel dans la passerelle domestique et interfacée avec les proxy de services,
- la définition des règles d'utilisation d'UPnP QoS v3 dans un environnement HGI,
- un algorithme d'admission d'appel réparti entre les passerelles domestiques pour gérer les ressources en sortie du DSLAM,

7.1. La passerelle résidentielle

La gestion du point de congestion sur la ligne de cuivre peut être confiée à la passerelle domestique. Elle matérialise la frontière entre le réseau domestique et le réseau de collecte et donc voit passer tous les flux entrant et sortant de l'utilisateur. Comme le réseau local est en adressage privé, la passerelle effectue une translation des adresses IP privées vers une adresse IP publique (fonction NAT). Ce changement d'adresse au niveau du plan de transfert impose dans certains cas une mise à jour des données de services. Ce rôle est dévolu au **proxy** ou au « Application Layer Gateway - **ALG** ». Il est alors possible, à partir de ces passerelles de service, d'extraire des messages des applications, les informations nécessaires à initialiser la requête de QoS auprès du module d'admission d'appel (voir figure 22). Une fois le flux admis, la QoS est configurée au niveau des interfaces de la passerelle domestique si nécessaire, puis le proxy ou l'ALG est averti que le flux est bien accepté. La signalisation applicative peut alors continuer son processus habituel. Par exemple, il est possible de déduire la bande passante du type de codec spécifié dans les messages SIP ou RTSP. Lorsqu'il n'est pas possible de trouver des informations dans la signalisation de l'application, il est alors nécessaire de formater la requête de QoS avec des données de service qui auront été approvisionnées dans la passerelle domestique e.g. le débit des chaînes TV.

Cependant, ce lien indirect avec les applications peut devenir un obstacle au déploiement de nouveaux services car il implique le déploiement d'un proxy ou ALG à chaque fois que le protocole applicatif change. Aussi, il est plus pertinent et pérenne de munir les terminaux d'un dispositif de communication générique, comme le standard UPnP QoS, avec la passerelle domestique afin de rationaliser les demandes de ressources préalablement à la demande de service. En cas de refus, la fonction d'admission d'appel notifie le demandeur (proxy, ALG ou terminal) qui a pour rôle de notifier en retour l'application et stopper la demande de service. Si la requête est acceptée, la fonction d'admission d'appel configure si

nécessaire la QoS appropriée (classification, police, marquage, mise en forme et file d'attente) dans la passerelle domestique avant de notifier le proxy ou le terminal demandeur qui poursuivra la demande de service via la signalisation applicative.



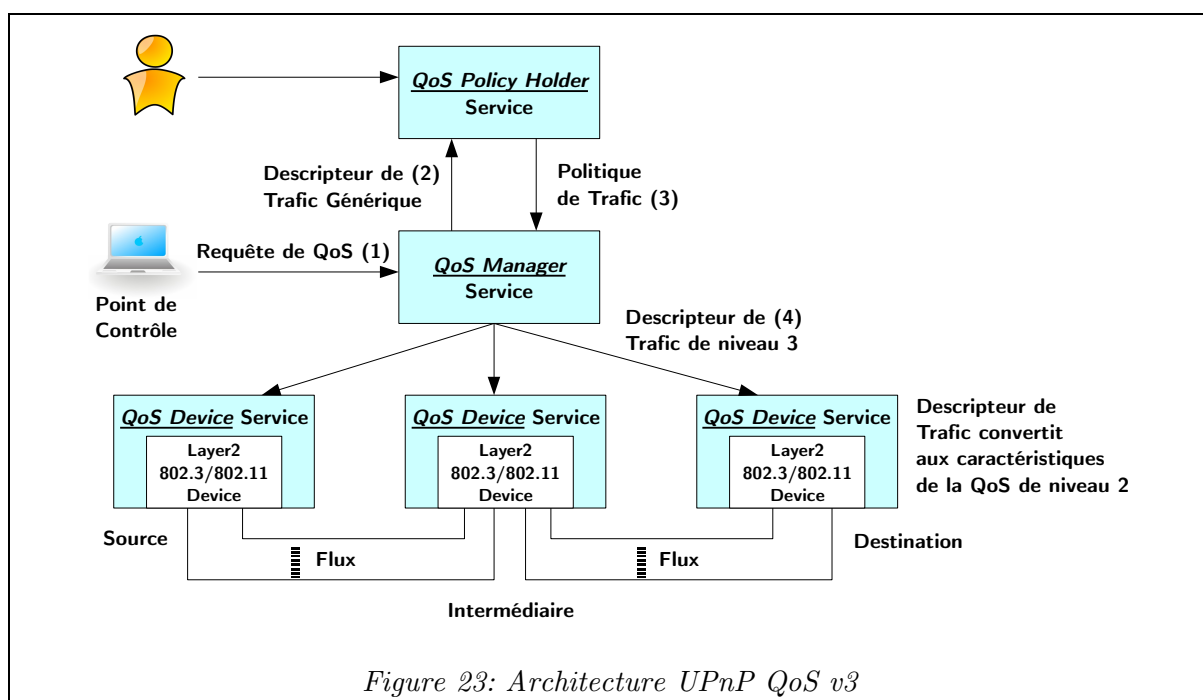
7.1.1. Compatibilité avec les standards

Le standard le plus avancé pour la gestion de la QoS dans le réseau domestique est la spécification produite par le forum Home Gateway Initiative – HGI au travers de ses spécifications [142]. Il définit essentiellement le contrôle de qualité de service par gestion différenciée des classes de services, mais spécifie également un mécanisme d'admission d'appel pour le service de voix sur IP. Le modèle proposé reste compatible avec la spécification HGI du fait qu'il a servi de base à la contribution [143] ayant abouti à la spécification de ce mécanisme d'admission d'appel. L'usage de la CAC est simplement étendu à l'ensemble des services du réseau domestique et non cantonné au seul service de voix sur IP. De plus, la spécification HGI introduit la notion de flux **gérés** pour lesquels les opérateurs s'engagent à garantir un niveau de QoS donné et de flux **non gérés** pour lesquels les opérateurs n'ont aucun engagement particulier à respecter. Le premier cas concerne les applications fournies par les opérateurs y compris un flux interne au réseau domestique mais issu d'un service de l'opérateur e.g. Une vidéo, diffusée depuis un serveur multimédia sur la télévision, sera considérée comme un flux géré, si elle a été téléchargée précédemment depuis le service de vidéo à la demande de l'opérateur.

Le deuxième standard pour la gestion de la QoS a été défini par le forum « Universal Plug and Play (UPnP) » au travers du service QoS [116]. Le service a évolué depuis la simple

gestion de priorité statique (version 1), puis dynamique (version 2) à l'allocation dynamique de ressources (version 3). Il est en réalité composé de 3 services (voir figure 23).

- Le « QoS Devices », implémenté dans chaque équipement et terminal du réseau domestique, est la base de l'architecture UPnP QoS. En plus de la gestion de la QoS différenciée par classe de service, il implémente la fonction d'admission d'appel sur ses interfaces.
- Le « QoS Manager » gère les demandes de QoS émanant des applications. Il est en charge de découvrir puis d'interroger tous les « QoS Devices » qui contribuent à l'acheminement des données entre la source et la destination.
- Enfin, le « QoS Policy Holder » fournit les règles à appliquer si la demande de QoS est du type priorité simple i.e. sans admission d'appel, et si il est nécessaire de préempter un service au bénéfice d'une application plus prioritaire.



Outre le caractère très verbeux du protocole SOAP utilisé dans la spécification UPnP QoS, la difficulté d'utilisation de ces services réside dans la possibilité de disposer de plusieurs « QoS Manager » et « QoS Policy Holder » dans le réseau domestique. Même si cela a été prévu, il empêche un opérateur de s'assurer que ce sont bien ses règles et ses politiques qui seront appliquées à ses services. L'exemple le plus probant est la préemption d'un service de l'opérateur, une communication téléphonique par exemple, par un flux du réseau domestique, par exemple un transfert de fichier. Il est impossible dans ces conditions de pouvoir garantir un niveau de qualité de service pour les flux gérés.

L'architecture proposée se voulant compatible à la fois avec les deux standards précédents, a nécessité de restreindre les cas d'usages des services UPnP QoS :

- La passerelle domestique doit implémenter le service « QoS Device » et si possible les services « QoS Manager » et « QoS Policy Holder »
- Les services « QoS Manager » et « QoS Policy Holder » doivent s'identifier comme

étant compatibles avec HGI,

- Le « QoS Policy Holder » doit contenir des règles et des politiques¹⁶ compatibles avec HGI notamment en supportant la notion de flux gérés et non gérés,
- La préemption ne sera supportée par le service « QoS Device » de la passerelle domestique que si elle émane d'un « QoS Policy Holder » via un « QoS Manager » identifié comme compatible avec HGI i.e. La vérification est simplifiée lorsque tous les services sont présents dans la passerelle,
- La passerelle n'acceptera de réserver des ressources dans la catégorie des flux gérés que si la demande émane d'un « QoS Manager » compatible avec HGI i.e. De même, l'identification sera aisée si tous les services sont présents dans la passerelle domestique,
- Enfin, les terminaux, supportant les services de l'opérateur pour lesquels un engagement de qualité de service est garanti, doivent spécifier l'utilisation du « QoS Manager » et du « QoS Policy Holder » compatible avec HGI lorsqu'ils émettent une requête de QoS.

Ces six règles recadrent l'utilisation d'UPnP QoS v3 afin de le rendre compatible avec la spécification HGI.

7.2. Admission d'appel répartie

La gestion du point de congestion pour l'accès au réseau de collecte peut être résolu par l'utilisation d'une fonction de type RACS telle que prévue dans l'architecture IMS. Cependant, cette solution présente plusieurs inconvénients dont l'acquisition de la topologie et l'intégration avec les services. La dernière contribution de ce mémoire vise à s'affranchir d'une plate-forme centralisée d'admission d'appel dans le réseau de collecte tout en offrant une solution à la gestion du point de contention en sortie du DSLAM.

Le principe proposé par l'auteur [144] est un nouveau mécanisme permettant de déployer une fonction d'admission d'appel localisée uniquement dans les passerelles domestiques et gérant à la fois la contention du lien d'accès à l'équipement de raccordement au réseau d'accès (DSLAM) ainsi que le lien de sortie de cet équipement (DSLAM) vers le réseau de collecte aussi bien sur le sens montant que descendant. Il repose sur la diffusion, via un arbre multicast, de l'information de disponibilité des ressources réseaux de l'équipement d'accès (DSLAM) vers les passerelles domestiques afin qu'elles puissent effectuer une admission d'appel local à la fois sur leur lien de raccordement à l'équipement d'accès au réseau et sur le lien de raccordement de l'équipement d'accès au réseau de collecte. Le système proposé peut être facilement déployé car il réside uniquement dans la passerelle résidentielle que l'opérateur maîtrise. Il ne nécessite pas d'ajout de plate-forme supplémentaire (de type CAC centralisée) ni de mise à jour des équipements réseaux (signalisation RSVP ou NSIS, sonde de mesures). Il est robuste et s'adapte aux pannes. Le passage à l'échelle est facilité car le système est borné par le nombre d'abonnés xDSL présent derrière un même DSLAM. De plus, il permet une bonne optimisation de l'utilisation des lignes ADSL et du lien de raccordement du DSLAM au réseau d'accès.

7.2.1. Détails du principe

Soit T_m le temps de propagation maximale des messages sur l'arbre multicast tel que :

$$T_m = T_{pu} + T_{pd} + T_r$$

¹⁶ La spécification UPnP QoS ne définit pas les politiques et règles contenues dans le « QoS Policy Holder » mais seulement le moyen d'y accéder et le format qu'elles doivent suivre.

avec T_{pu} le temps de propagation upstream, T_{pd} le temps de propagation downstream et T_r le temps de réplication du paquet IP multicast dans le DSLAM.

Sur une liaison xDSL, $T_{pu} = T_{pd} \gg T_r$ soit $T_m = 2 * T_p \approx 160$ ms

Les débits contrôlés sont :

- $\Lambda_{G_{phy}}$: Le débit physique du lien de raccordement entre l'équipement d'accès et le réseau de collecte
- $\Lambda_{G_{CoS}}(c)$: Le débit de la Classe de Service c entre l'équipement d'accès et le réseau de collecte
- $\Lambda_{L_{phy}}$: Le débit physique du lien de raccordement entre la passerelle domestique et l'équipement d'accès
- $\Lambda_{L_{CoS}}(c)$: Le débit de la Classe de Service c entre la passerelle domestique et l'équipement d'accès

La configuration initiale du dispositif est la suivante (voir figure 24) :

- Le DSLAM sur lequel sont raccordées les passerelles domestiques est configuré avec une adresse multicast dédiée à la fonction de CAC. Cette adresse multicast a la particularité de correspondre à un arbre multicast ne comportant qu'un seul point de réplication i.e. le DSLAM lui-même et n feuilles i.e. les passerelles domestiques. De plus, cet arbre multicast est configuré de telle sorte que chaque passerelle domestique i.e. chaque feuille est autorisée à émettre des paquets multicast. Ainsi configuré, l'arbre multicast se comporte comme un bus sur lequel chaque entité peut émettre un message à destination de toutes les entités déjà connectées.
- Chaque passerelle domestique est configurée avec l'adresse multicast dédiée à la fonction d'admission d'appel ainsi que les valeurs nominales de débit de chaque classe de service supportée par le DSLAM et la valeur nominale de l'interface physique de raccordement du DSLAM au réseau de collecte.
- Au temps t_0 , lorsqu'une passerelle se connecte au DSLAM, la passerelle émet sur l'adresse multicast une trame HELLO destinée à se faire connaître des autres passerelles et récupérer l'état en cours des réservations via des messages UPDATE.
- Chaque passerelle gère en local la disponibilité des ressources sur son lien de raccordement vers l'équipement d'accès ainsi qu'une copie des ressources du lien de sortie de l'équipement d'accès.

Lorsqu'une passerelle veut réserver des ressources, elle vérifie la disponibilité des ressources dans la classe de service demandée et sur son lien de raccordement à l'équipement d'accès, de sorte à ce que la somme des débits des applications locales n'excède pas les limites de la ligne de cuivre $\Lambda_{L_{CoS}}(c)$ et $\Lambda_{L_{phy}}$. Puis, elle vérifie que les ressources sont également disponibles sur le lien de sortie de l'équipement d'accès via sa copie locale, de sorte à ce que la somme des débits de toutes les passerelles domestiques n'excède pas les limites de l'interface du DSLAM $\Lambda_{G_{CoS}}(c)$ et $\Lambda_{G_{phy}}$.

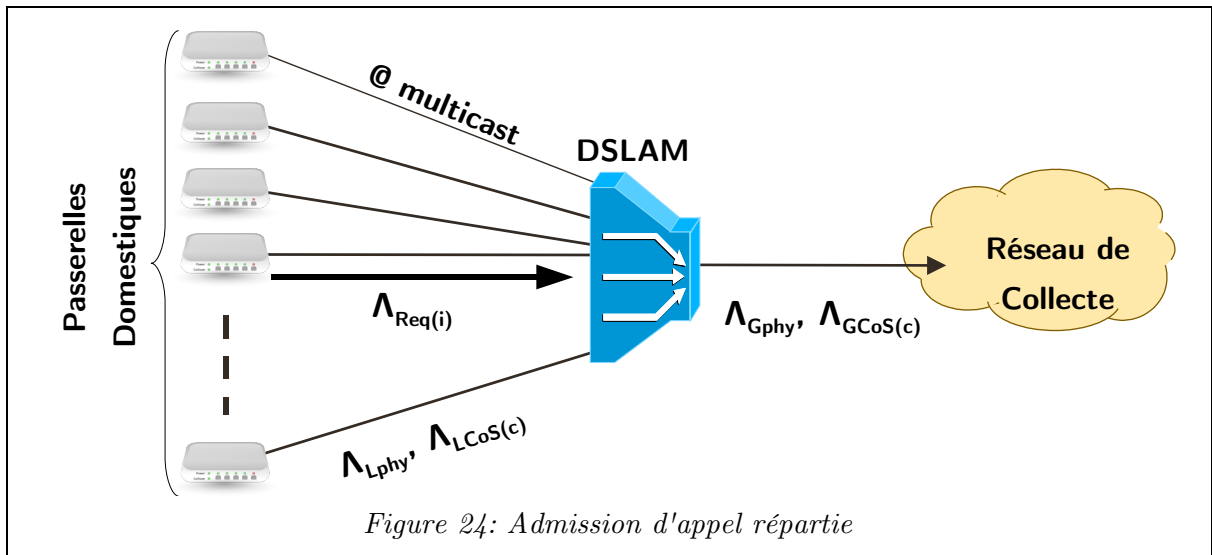
Si les deux conditions sont remplies, elle émet un message REQUEST sur l'arbre multicast contenant sa réservation. Toutes les autres passerelles vont alors recevoir ce message et mettre à jour leur copie locale et en particulier la passerelle primaire. Au bout d'un temps $2 * T_m$ nécessaire à la propagation de la requête et sa prise en compte, la passerelle confirme la réservation en émettant un message CONFIRM avec les nouvelles valeurs de débit qu'elle consomme.

Si deux ou plus passerelles émettent simultanément un message REQUEST, une collision sera détectée avant un temps égale à $2 * T_m$. Si les ressources ne sont pas suffisantes,

toutes les requêtes sont alors annulées. Chaque passerelle attend un temps aléatoire puis re-émet sa requête.

Chaque passerelle émet régulièrement, à intervalle T_u , un message UPDATE pour notifier la quantité de ressources qu'elle continue d'utiliser. L'ensemble des messages UPDATE est consolidé par chaque passerelle domestique afin de mettre à jour les ressources consommées par l'ensemble du groupe. Ainsi, l'ensemble des passerelles se synchronise tous les T intervalles. Si une passerelle n'émet pas un message UPDATE avant la période T_u , en cas de panne ou de problème, ses réservations sont automatiquement retirées de la somme des ressources consommées et elle est retirée de la liste des passerelles consommatrices. La passerelle en défaut ne pourra de nouveau réserver de la ressource que si elle émet au préalable un message HELLO.

Une fois la réservation de débit arrivée à son terme, la passerelle domestique met à jour la copie locale de ses réservations et envoie un message RELEASE avec les débits qu'elle libère. Lorsqu'une passerelle ne consomme plus aucune ressource, elle envoie un dernier message UPDATE avec les valeurs de débit positionné à 0 puis cesse d'émettre des messages UPDATE. Les autres passerelles la retirent de la liste des passerelles à surveiller. Cette dernière redeviendra active après envoi d'un message REQUEST.



7.2.2. Applications

Couplé avec la fonction d'admission d'appel et de réservation de ressources située dans la passerelle domestique et contrôlant les ressources du réseau domestique, il est possible de vérifier la disponibilité du débit non seulement sur la ligne de cuivre, mais également en sortie du nœud de raccordement au réseau de collecte. La QoS est donc garantie par réservation des ressources et le marquage des paquets pour qu'ils soient aiguillés vers la bonne classe de service.

IV

Bilan et Perspectives

« One policy, one system, universal service »

T. Vail, premier président d'ATT, 1907

Cette dernière partie du mémoire est consacrée au bilan des travaux sur la gestion de la QoS et, à partir de ce dernier, propose une solution de gestion de bout en bout ainsi que les travaux prospectifs s'y rapportant.

D'abord, la synthèse des travaux nous permet de définir des règles d'urbanisme à respecter en matière de gestion de la QoS. Ces règles permettent de sélectionner le meilleur mécanisme pour la bonne partition de réseau et au bon instant.

Une proposition d'architecture de référence, respectant ces règles, est ensuite présentée.

Finalement, la dernière section aborde les perspectives des travaux restant à couvrir, en particulier dans le cadre de cette architecture de référence.

1. Bilan des travaux

La dernière partie de ce mémoire propose une solution globale tirant partie des avantages des différentes solutions collectées au fil du temps ainsi qu'un modèle d'architecture répondant à la problématique du contrôle et de la garantie de la qualité de service de bout en bout sur des réseaux hétérogènes. Elle est issue des travaux menés par l'auteur dans ce domaine depuis une dizaine d'années de façon cohérente et dont les fondements principaux sont :

- Une gestion des files d'attentes et du multiplexage par espacement comme discipline de service universel garantissant un délai, une gigue et un taux de perte faible tout en respectant la caractéristique temporelle intrinsèque de chaque application,
- Une gestion par Classes de Services approvisionnées par MPLS-DiffServ-TE à l'intra et à l'inter domaine durant le processus de commande / livraison,
- Une admission d'appel pour protéger les ressources affectées aux classes de services durant le processus d'invocation,
- Une surveillance et une mesure de la QoS fournies aux utilisateurs pour vérifier les engagements des opérateurs.

À partir de la synthèse des travaux, des règles d'urbanisme ont été élaborées afin de concevoir l'architecture proposée.

1.1. Synthèse

Les travaux de recherches présentés dans ce mémoire (voir figure 25) ont tout d'abord débuté par l'étude du multiplexage par espacement des cellules ATM. La technique d'espacement initialement appliquée pour la conformité et la mise en forme des flux, s'est révélée être une véritable discipline de file d'attente. Elle permet, à l'aide d'un seul algorithme, de gérer les contraintes de plusieurs classes de services en conservant la caractéristique temporelle des flux. Ainsi, au sein de la même file d'attente, l'algorithme du VSA est capable de promouvoir une cellule vers la sortie même si elle est arrivée bien après les cellules des autres connexions. L'espacement permet également de réguler sans perte les flux TCP dès lors que la file d'attente est capable de stocker une fenêtre TCP. En protégeant le réseau et les connexions, cet algorithme garanti une QoS absolue à l'ensemble des connexions traitées. De même, TCP s'accommodant de variation de débit dès lors qu'il est régulé par espacement, il est intéressant de moduler le débit de l'espaceur en fonction des ressources disponibles dans le réseau. La signalisation allégée, mise au point dans le projet ASIA, a permis de coupler TCP avec les ATC ABT/DT et ABT/IT pour réserver du débit au sein de tunnels MPLS afin de garantir de la QoS tout en optimisant le taux de remplissage du réseau.

Cependant, le couplage de TCP avec la signalisation ABT est limité et ne permet pas de déterminer tous les paramètres nécessaires à la gestion de la QoS. Aussi, l'architecture à base de médiation de service et médiation de ressource, proposée dans le projet CADENUS, facilite la négociation de la QoS et la configuration automatique du réseau par translation des SLA en SLS. La QoS est garantie à la fois par la fonction d'admission d'appel et la configuration ad-hoc des routeurs du réseau. Elle a ensuite été complétée par la définition de l'interface entre la médiation de service et la médiation de réseau afin d'offrir l'accès à un service de QoS à la demande accessible depuis les plates-formes de service ou les applications.

Les prémices des relations entre opérateurs ébauchées dans l'architecture CADENUS ont ensuite été approfondies au travers de l'architecture, à base d'approvisionnement de

chemins à QoS, proposée dans le projet EuQoS. Ces chemins à QoS sont matérialisés par l'emploi de tunnels MPLS-DiffServ-TE qui réservent des ressources par classes de services. Les serveurs PCE coopèrent entre-eux à l'inter-domaine afin de calculer le placement optimal des tunnels à travers les différents opérateurs de réseaux impliqués. Les tunnels MPLS-DiffServ-TE sont ensuite protégés par une fonction d'admission d'appel localisée dans les gestionnaires de ressources RM de chaque domaine. L'ensemble permet de garantir une QoS absolue pour les connexions empruntant le système EuQoS.

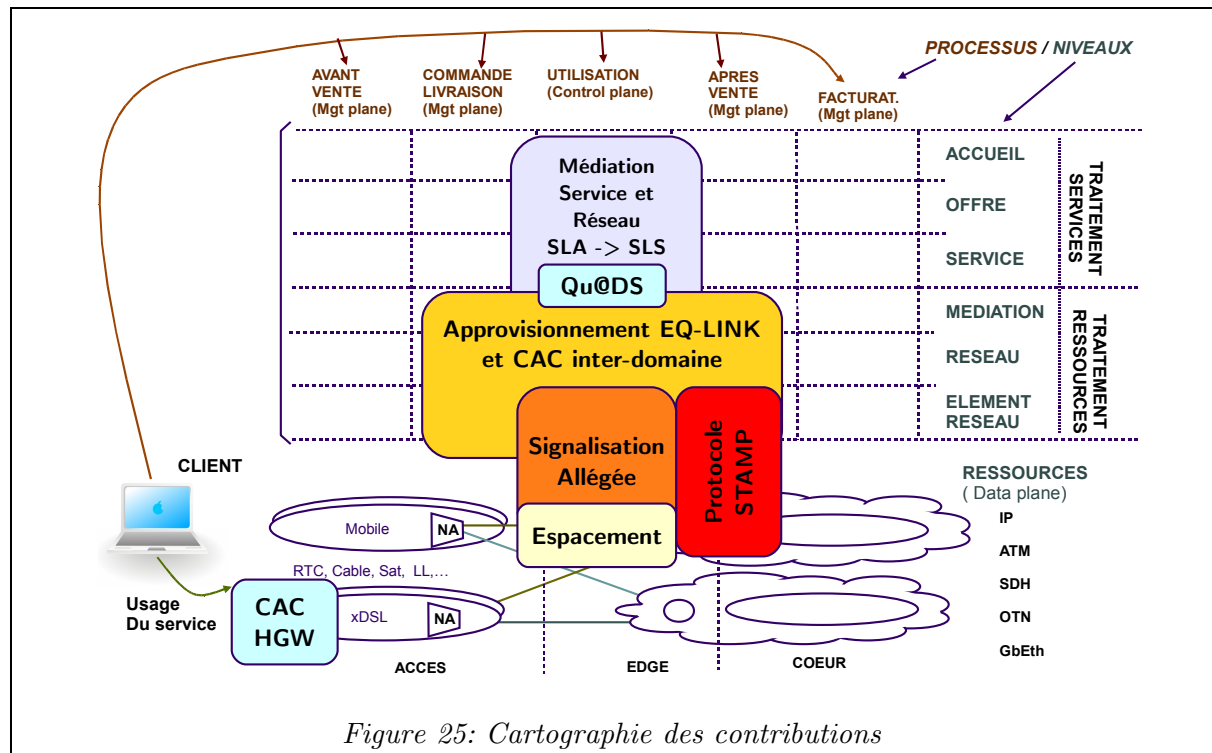


Figure 25: Cartographie des contributions

L'admission d'appel nécessite cependant une parfaite connaissance du réseau qu'elle contrôle. Un nouveau protocole d'acquisition de la topologie du réseau a ainsi été élaboré pour alimenter la base de données de la CAC. La volumétrie des informations a ensuite poussé à agréger les données et développer de nouveaux schémas d'admission d'appel sur ces topologies agrégées.

Finalement, les travaux ont porté sur la gestion de la QoS dans le réseau domestique. En particulier, l'implémentation d'une fonction de CAC dans la passerelle domestique permet de vérifier la disponibilité des ressources dans le réseau local et dans le réseau d'accès. La passerelle permet également de simplifier le lien entre les applications et le réseau pour les demandes de QoS.

1.2. Règles d'urbanisme

Un transfert de données entre une source et une destination donnée va nécessairement traverser plusieurs réseaux. Soit P_r , la « profondeur du réseau », la distance entre la source ou la destination et le point d'observation du réseau où l'on se situe. Une profondeur de « 1 » correspondra à la distance minimale observable dans le premier ou le dernier réseau impliqué dans le transfert e.g. le réseau domestique. La profondeur sera maximale lorsque le point d'observation sera situé au milieu du gué i.e. lorsque l'on se situe à mi-chemin entre la source

et la destination en terme de réseaux traversés e.g. dans un réseau de transit.

Les règles d'urbanisme suivantes sont basées sur le constat que les mécanismes de gestion de la QoS doivent migrer progressivement du processus d'invocation vers le processus de commande/livraison au fur et à mesure de la progression dans le réseau, lorsque P_r augmente.

Règle 1 : *Le mécanisme de QoS doit s'appliquer aux trafics proportionnellement agrégés à la profondeur du réseau P_r où il s'applique.*

Ainsi, un mécanisme de QoS proche de l'émetteur ou du récepteur doit contrôler le flux ou le micro flux, alors que disposé dans un réseau de transit il ne doit s'occuper que des agrégats de haut niveau (e.g. Agrégation du trafic de la classe EF en provenance de l'AS A et à destination de l'AS B). Le contrôle session par session doit s'appliquer en périphérie de réseau et le contrôle par priorité et ingénierie de trafic dans les réseaux de cœurs et de transits. Cette première définition tire ses origines de deux constats :

- Il n'est pas possible de contrôler finement les flux ou les micro flux dans un réseau de transit du fait du passage à l'échelle. En effet, les équipements devraient pouvoir gérer plusieurs dizaines voire plusieurs centaines de millions de contextes et identifier finement les flux et micro flux.
- Il est nécessaire de pouvoir vérifier au plus tôt la conformité des flux et les mettre en forme afin de faciliter leur multiplexage dans le réseau et pour protéger ce dernier ainsi que les flux déjà admis.

Règle 2 : *Le mécanisme de QoS doit s'appliquer à une échelle de temps inversement proportionnelle à la profondeur du réseau P_r .*

Pour des raisons similaires à la première définition, les mécanismes de contrôle de la QoS doivent s'appliquer à la même échelle de temps que les trafics qu'ils gèrent. Dans le cas des réseaux d'accès où sont traités les flux et micro flux, il faut pouvoir appréhender l'ensemble des demandes de services en provenance des clients. A l'inverse, dans le réseau de transit, et encore une fois pour des raisons de passage à l'échelle, il faut pouvoir appréhender seulement les demandes de traitement des agrégats de trafics qui émanent généralement des opérateurs eux-mêmes ou de leur plus gros clients.

Règle 3 : *Le mécanisme de QoS doit configurer les équipements du réseau à une fréquence inversement proportionnelle à la profondeur du réseau P_r .*

Corollaire des deux premières définitions, le mécanisme de contrôle de QoS doit configurer session par session les équipements de périphérie de réseau afin d'assurer les fonctions de classification, police, marquage et mise en forme des flux. Dans les réseaux cœur ou de transit, où ne sont traités que les agrégats, la configuration des équipements s'effectue uniquement durant l'approvisionnement du réseau (qu'il soit statique ou dynamique).

Règle 4 : *Une demande de QoS est inversement liée au service en fonction de la profondeur de réseau P_r .*

Cette définition traduit le lien existant entre les applications et la gestion session par

session ou par agrégat. Dans le cas des sessions, la demande de QoS doit émaner du service (application cliente ou serveur) ou du proxy de service afin de pouvoir déterminer avec fiabilité le début, la fin et le débit ainsi que la classe de service requise par la session. Inversement, la gestion des agrégats résulte d'une mise en place par l'ingénierie de trafic qui, certes dépend de la classe de service, mais est indépendante d'une application donnée.

Règle 5 : *Un contrôle de QoS est toujours lié à la classe de service auquel appartient l'application*

Cette définition impose de gérer la QoS par classe de service et non globalement. En effet, la sensibilité différente des applications aux paramètres de QoS que sont le débit, le délai, la gigue et la perte, ne permet pas de traiter toutes les applications avec les mêmes contraintes. Il faut donc pouvoir disposer de plusieurs classes de services qui regroupent les applications sensibles aux mêmes paramètres de QoS e.g. Sensitif au délai et la gigue, sensible à la perte ... Enfin, une dernière classe de service dite « Best Effort » ou au mieux sera systématiquement implémentée afin de transférer toutes les applications qui ne requièrent pas de QoS. En effet, toutes les applications ne nécessitent pas de QoS. Cependant il faut pouvoir les isoler à part afin qu'elles ne perturbent pas les applications qui requièrent de la QoS.

Règle 6 : *Une fonction d'admission d'appel et de réservation de ressources doit s'opérer sur une vue synchrone des éléments qu'elle contrôle*

Dernière règle, elle impose que la topologie du réseau doit être parfaitement synchrone avec sa représentation dans la base de données de la fonction d'admission d'appel. Si la topologie ou le routage change sans que la fonction de CAC soit mise à jour, les décisions de l'algorithme pour les requêtes suivantes risquent d'être fausses. À la fois parce que le chemin emprunté par les flux ne correspond pas à celui qui est contrôlé, mais surtout parce que les réservations en cours ne seront pas mises à jour et donc ne reflèteront pas l'usage réel des liens contrôlés.

En résumé :

- Les mécanismes de contrôle de QoS doivent s'appliquer sur le mode session par session au plus près de la source,
- Les requêtes de demandes de ressources vers la fonction d'admission d'appel doivent émaner des applications,
- La gestion de la QoS par classes de services différenciées est ensuite utilisée dans le reste du réseau.

1.3. Architecture de référence

L'architecture générique que nous proposons (voir figure 26) est basée sur les règles d'urbanisme énoncées précédemment et tire partie du meilleur du contrôle de la QoS session par session et par agrégat en cherchant **quand** les mécanismes de gestion de QoS doivent intervenir, entre le processus de commande/livraison et le processus d'invocation et **où** il doivent s'appliquer en fonction des partitions du réseau.

Le principe générique repose sur une gestion de la QoS dans 4 classes de services : Conversationnel, Multimédia, Données et « Best Effort ». Chaque classe de service se voit attribuer des caractéristiques de QoS appropriées à chaque type de service :

- Classe **Voix** : délai court, gigue faible et peu de perte pour le Conversationnel,
- Classe **Vidéo** : délai faible, gigue constante, pas de perte pour le Multimédia,
- Classe **Donnée** : délai et gigue quelconque, pas de perte pour les Données,
- pas de contrainte pour le « **Best Effort** »

Chaque classe de service est approvisionnée en débit durant le processus de Commande/Livraison par ingénierie de trafic. Durant le processus d'invocation, les ressources de chaque classe (essentiellement le débit) sont allouées par une fonction d'admission d'appel en entrée du réseau qui protège ainsi les classes de services contre la surcharge. La QoS est finalement mesurée et surveillée dans chaque classe de service durant le processus d'après-vente pour vérifier que les engagements contractuels sont bien respectés et agir en cas de défaillance.

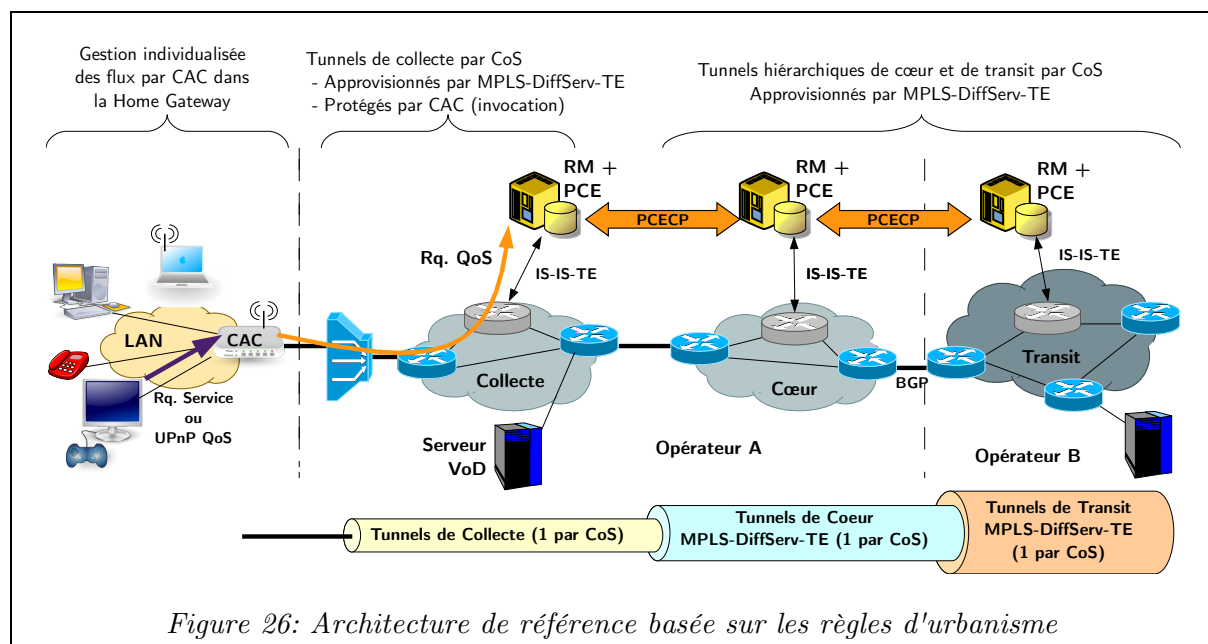


Figure 26: Architecture de référence basée sur les règles d'urbanisme

Ainsi, le réseau local sera géré à l'aide d'une fonction d'admission d'appel opérant dynamiquement session par session avec configuration dynamique des équipements afin de classer, policer, marquer et mettre en forme le trafic en fonction de la QoS et la classe de service demandée. Ces dernières sont implémentées dans les réseaux suivants à l'aide de tunnels MPLS-DiffServ-TE : un tunnel par classe de service pour une destination (préfixe réseau) donnée. La hiérarchisation des tunnels à chaque changement de réseau permettant enfin d'étendre la solution à grande échelle.

1.3.1. Réseaux locaux

L'architecture proposée pour les réseaux domestiques ou des petites entreprises (SOHO) est centralisée autour de la passerelle domestique telle que présentée au chapitre §III.7. Cette dernière embarque une fonction d'admission d'appel associée aux classes de services à même de gérer les ressources du réseau domestique. Au vu du taux très faible de demandes (maximum 10 par heure), il est possible de configurer dynamiquement la passerelle domestique sans impacter les performances des services afin de classer, policer, marquer et mettre en forme les flux session par session en utilisant le multiplexage par espacement. De même, le faible nombre de sessions simultanées autorise la fonction d'admission d'appel à

gérer les états – procéder « state-full » - sans impact sur les performances. L'interaction avec les services est assurée implicitement par les proxy embarqués dans la passerelle ou explicitement en utilisant UPnP QoS v3.

La passerelle domestique étant gérée et authentifiée par l'opérateur, elle est tiers de confiance vis à vis des flux émis vers le réseau longue distance autorisant, de ce fait, les requêtes de QoS vers le réseau. Une fois le flux admis, elle va pouvoir appliquer le marquage approprié. Le réseau d'accès pourra s'appuyer sur ce marquage pour classer et gérer les flux non session par session, mais par agrégat au sein d'une même classe de service. Finalement, la fonction d'admission d'appel s'exerce, non seulement sur les liens et équipements du réseau domestique, mais également sur le lien de raccordement au réseau d'accès, en particulier pour la technologie ADSL. La passerelle domestique délivre ainsi des flux conformes à la politique de l'opérateur et vérifie qu'il n'y a pas de congestion potentielle ni sur le lien de raccordement ni dans le réseau domestique lui-même.

Dans le cas des grands réseaux, la passerelle domestique est remplacée par un routeur dédié appelé « Customer Premise Equipment » (CPE) installé en face du point de raccordement au réseau d'accès matérialisé par un routeur appelé « Provider Edge » (PE). Le CPE va de nouveau fournir les mêmes fonctionnalités que la passerelle domestique. Cependant, si la gestion des priorités par classes de services est maintenue, la fonction d'admission d'appel est souvent implicite. Le nombre de postes clients, donc de sessions simultanées, dans un grand réseau d'entreprise ne favorise pas l'utilisation de la CAC. Ainsi, les classes de services sont généralement implémentées via des VLAN sur les commutateurs Ethernet constituant le réseau local. Les flux ou terminaux sont alors aiguillés dans le bon VLAN lors de la configuration ou du raccordement. Le marquage des flux est réalisé implicitement en fonction du VLAN. Basé sur un système de liaisons louées plutôt que sur un système de connexions à la demande, les ressources nécessaires sur le lien de raccordement sont approvisionnées en fonction du besoin global de l'entreprise. Un lien avec le système d'approvisionnement des ressources de l'opérateur permet au client de redimensionner ses connexions en fonction de ses besoins.

Les flux émanant des plates-formes de services de l'opérateur sont également gérés via le nœud de raccordement au réseau. Cet équipement tient le même rôle que la passerelle domestique ou le CPE. Il identifie, police, marque et met en forme les flux de services. L'identification des applications est réalisé simplement car l'opérateur maîtrise à la fois le plan d'adressage (utilisation de l'adresse IP source) et les services offerts (utilisation du numéro de port source).

1.3.2. Réseaux de collectes

L'approvisionnement des ressources, pour chaque classe de service supportée, peut prendre plusieurs formes selon la technologie de transferts : VP en ATM, VLAN pour le GigaBit Ethernet, ... Lorsque c'est possible, la création de tunnels MPLS-DiffServ-TE sera privilégiée. La gestion des ressources physiques du réseau sera confiée au serveur de placement de tunnel PCE. La souplesse dans la gestion par ingénierie de trafic, la possibilité de protéger les tunnels et l'homogénéisation sont autant d'atouts pour simplifier la tâche des opérateurs.

Aucune configuration d'équipement n'est nécessaire puisque le réseau local, via la passerelle résidentielle ou le CPE, a effectué l'admission du nouveau flux et ajusté le marquage en fonction de la classe de service demandée. Le marquage en couche IP sera privilégié afin d'être le plus indépendant possible du plan de transfert. Côté plate-forme de service, les points de raccordement des serveurs, i.e. les nœuds d'accès, sont configurés de

façon statique de sorte à marquer les flux sortant avec le bon code DSCP.

Afin de protéger l'utilisation des tunnels MPLS-DiffServ-TE, il est nécessaire de déployer une fonction d'admission d'appel pour contrôler précisément le trafic qu'ils doivent acheminer. Le déploiement d'une telle fonction peut s'effectuer par plaque ou point de présence (PoP) afin de réduire la volumétrie des requêtes et mieux supporter le passage à l'échelle par rapport à un déploiement sur un réseau complet. Deux méthodes sont alors possibles selon la configuration des tunnels :

- Si l'approvisionnement a tissé un maillage complet de tunnels entre les nœuds d'accès au réseau de collecte et les nœuds d'accès au réseau de cœur, l'algorithme d'admission d'appel réparti entre les passerelles domestiques (voir §III.7) peut être utilisé pour contrôler l'accès aux tunnels.
- Dans les autres cas de figure, il est alors nécessaire de déployer une fonction d'admission d'appel centralisée au niveau du réseau de collecte. Deux contraintes de mise en œuvre sont résolues dans ce cas d'application. L'acquisition et la maintenance de la topologie sont facilitées car il suffit d'écouter les protocoles de routage à ingénierie de trafic tel que IS-IS-TE ou OSPF-TE puisque les tunnels MPLS-DiffServ-TE y sont annoncés.

La protection des tunnels avec des tunnels de secours évite à la fois de devoir mettre à jour la fonction d'admission d'appel avec une nouvelle topologie et réduit le temps d'indisponibilité du réseau grâce à la commutation des paquets sur le tunnel de secours en moins de 50ms. En deuxième lieu, il est possible de tirer partie de la fonction d'admission d'appel localisée dans la passerelle domestique pour invoquer la fonction d'admission d'appel du réseau de collecte. Il n'est alors pas nécessaire de modifier les plates-formes de services pour qu'elles s'interfacent avec la fonction de CAC.

Les points de congestions dans le réseau de collecte étant limités (en général le DSLAM et le nœud de raccordement), il est également possible d'utiliser une signalisation de type RSVP ou NSIS pour vérifier la disponibilité des ressources entre la passerelle domestique et le réseau de collecte. Le nombre de messages de signalisation étant limité (2 à 3 messages x 1000 à 32000 passerelles par équipement), il n'y a pas de facteur limitatif pour les performances ou le nombre d'états à gérer au niveau du DSLAM ou du nœud de collecte.

1.3.3. Réseau cœur

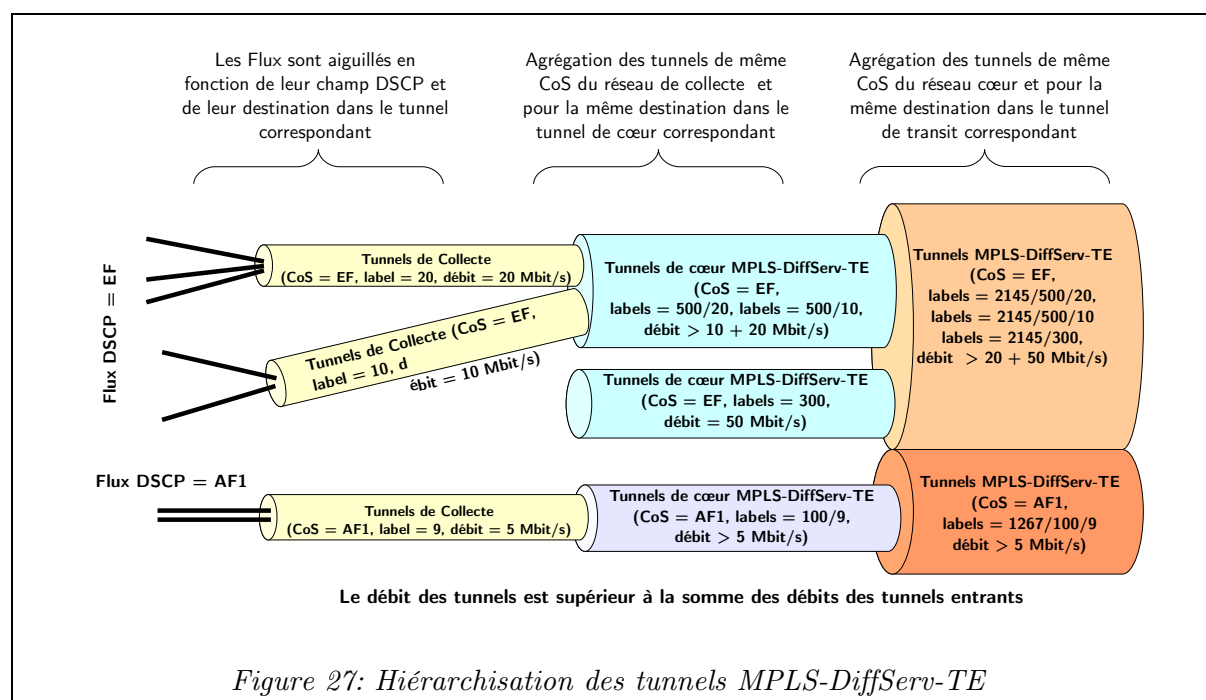
Les réseaux de cœur des opérateurs sont essentiellement des « backbone » IP constitués de routeurs. La signalisation MPLS-DiffServ-TE étant généralement disponible, l'architecture proposée est systématiquement basée sur l'approvisionnement de tunnels TE par classe de service. De façon similaire au réseau d'accès, les nœuds de bordure du réseau cœur sont maillés entre eux à l'aide de tunnels MPLS-DiffServ-TE. Ces derniers fournissent la QoS par classe de service avec une bande passante réservée. En suivant les règles d'urbanisme, la QoS doit être gérée au niveau des agrégats de flux et non session par session. Si dans le réseau de collecte il est encore possible de contrôler l'usage des tunnels via une fonction d'admission d'appel, ce n'est plus le cas pour les réseaux de cœur à cause du nombre trop élevé de connexions simultanées transitant dans le réseau cœur. Un deuxième niveau de label est utilisé en cœur de réseau afin d'agréger, par classe de service, les tunnels en provenance du réseau de collecte et pour la même destination dans les tunnels du réseau cœur. En effet, afin d'éviter l'emploi de l'admission d'appel dynamique session par session, c'est durant la phase d'approvisionnement que l'ingénierie de trafic contrôle le remplissage des tunnels du réseau cœur. Un tunnel du réseau cœur entre deux nœuds de bordure sera dimensionné avec un débit supérieur ou égal à la somme des débits des tunnels entrant dans ce dernier. De même,

les nœuds de bordure sont configurés, au moment de l'établissement des tunnels, avec des règles de transferts réduites à leur plus simple expression au sens MPLS du terme : le routeur empile les labels en fonction de leur point d'entrée (label et port physique) vers le point de sortie (label et port physique).

1.3.4. Inter Domaine

La gestion de la QoS à l'inter-domaine est basée sur les études menées dans le projet EuQoS et plus particulièrement sur le « hard model » [121]. La solution est similaire à celle proposée dans le réseau cœur. À nouveau, il s'agit d'établir des tunnels MPLS-DiffServ-TE avec le support du serveur PCE pour implémenter les différentes classes de services. La différence principale vient de l'introduction de l'algorithme de « Backward Recursive Path Computation (BRPC) » standardisé par le groupe de travail PCE de l'IETF. Cet algorithme permet aux PCE de coopérer afin de calculer des chemins qui s'établissent sur plusieurs domaines ou AS. Les opérateurs qui souhaitent fournir de la QoS à l'inter-domaine doivent nécessairement coopérer pour définir le cadre des échanges de trafic. Ces contrats sont établis de paire à paire de la même façon que pour BGP. A ceci près qu'il s'agit ici de définir les paramètres de QoS supportés et le volume de trafics échangés en plus des réseaux annoncés. Ces contrats sont initialisés dans chaque plateforme de gestion de ressources et servent à déclencher l'établissement des tunnels inter-domaines portant les classes de services pour le transit des flux à l'inter-domaine.

Les tunnels hiérarchiques (voir figure 27) seront de nouveau utilisés comme entre le réseau d'accès et le réseau cœur. Un troisième niveau de label est utilisé pour agréger les tunnels en provenance du réseau cœur vers des tunnels de transits inter-domaines qui acheminent tout le trafic vers le même AS de destination. La solution reste ainsi tout à fait extensible puisque le nombre de tunnels, y compris en transits, reste borné par le nombre de routeurs de bordures et le nombre de classes de services. Cela permet également de protéger plus efficacement les tunnels, puisqu'ils sont gérés indépendamment au sein de chaque domaine.



2. Perspectives

Le contrôle de la QoS reste un sujet vaste nécessitant encore des études avant de voir son déploiement se généraliser. En particulier, les efforts devront porter sur les relations à l'inter-domaine et la possibilité de choisir avec une plus grande flexibilité le chemin d'AS que doit suivre une connexion. Le routage hiérarchique à partir de l'agrégation de topologie semble une voie prometteuse et complémentaire de la technologie PCE pour l'établissement de tunnels MPLS-TE à l'inter-domaine. Le contrôle de la QoS session par session dans le réseau est également un facteur de succès pour les opérateurs. Le réseau domestique devenant de plus en plus complexe auprès d'une clientèle de plus en plus néophyte, les opérateurs se doivent de déployer les mécanismes les plus appropriés pour faciliter et garantir le bon fonctionnement de leurs services. Les enjeux majeurs résident dans la maîtrise de l'interaction de service incluant la notification et la préemption des applications en cas de problèmes et la mise au point de solutions robustes aux conditions sévères de l'environnement domestique. Cela nécessitera une normalisation du point d'accès au service réseau de gestion de la qualité de service afin d'être largement adoptée par les développeurs d'application. Enfin, le déploiement de la fibre optique à l'accès va peut-être accélérer le déploiement des fonctions d'admission d'appel dans la mesure où les accès deviennent symétriques et offrent aux clients un débit dans le sens montant 10 à 100 fois plus élevés qu'aujourd'hui. Ceci ouvre de nouvelles perspectives pour des applications de partage de contenus multimédia en temps réel nécessitant des débits élevés (de l'ordre de plusieurs dizaines de Mbit/s) pour lesquels une garantie de QoS est indispensable. Cette dernière section présente les principaux points d'études prospectifs qu'il sera nécessaire d'étudier dans les prochaines années, à savoir :

- La hiérarchisation des protocoles de routage à ingénierie de trafic pour calculer le chemin d'AS pour alimenter le PCE dans ses calculs de placement de tunnels,
- L'amélioration des algorithmes d'admission d'appel dans le réseau domestique pour tenir compte des technologies de transferts à débit fluctuant comme le WiFi ou le PLT,
- La préemption et l'adaptation des applications en fonction de la disponibilité des ressources et des priorités en applications,
- La mise en place automatique de QoS entre réseaux domestiques pour le partage de contenus multimédias autoproduits en temps réel.

Ces études correspondent aux pièces manquantes dans le schéma d'architecture présenté précédemment et dont les axes de recherches sont détaillés ci-après.

2.1. Le routage hiérarchique

2.1.1. Calcul du chemin d'AS par CoS

La mise en place de la QoS dans les réseaux cœurs et à l'inter-domaine doit s'opérer par ingénierie de trafic durant la phase approvisionnement des ressources du réseau. Des tunnels à classes de services différenciés sont établis à l'aide des outils DiffServ-MPLS-TE et du serveur de calcul de placement des tunnels PCE. Ces deux technologies reposent sur l'utilisation de protocoles de routage véhiculant des informations d'ingénierie de trafic : IS-IS-TE ou OSPF-TE. Ces deux protocoles, même s'ils permettent une hiérarchisation du réseau par découpe en aire ou région, restent dédiés à l'intra-domaine et ne sont jamais utilisés entre opérateur. Lors de l'utilisation de l'algorithme BRPC, les serveurs PCE ne peuvent tirer partie que des informations d'ingénierie de trafic disponibles dans leurs domaines respectifs et sont totalement aveugles vis à vis de l'inter-domaine.

Bien qu'ayant défini dans le protocole PCEP la possibilité de spécifier un chemin d'AS pour alimenter l'algorithme de BRPC, le groupe de travail PCE a exclu de son champ de normalisation le calcul de ce chemin d'AS. Les solutions pour fournir le chemin d'AS résultent d'un choix basé sur BGP. Pour les raisons évoquées dans les chapitres précédents, ce chemin suit le chemin dit « Best-Effort » et surtout ne donne pas de visibilité sur des chemins alternatifs qu'il serait aisé d'emprunter surtout pour l'établissement de tunnels. Plusieurs solutions ont été envisagées pour pallier ce problème. Tout d'abord, l'utilisation de versions modifiées de BGP tel qBGP ou EQ-BGP permettent de véhiculer des informations sur la QoS disponibles dans les différents AS et donc de déterminer un chemin d'AS différent en fonction de la classe de service ou des exigences de QoS.

2.1.2. Concept de hiérarchisation

PNNI permet de hiérarchiser la topologie afin de rendre l'architecture non seulement extensible aux réseaux de grande taille mais également de pouvoir gérer les demandes entre opérateurs sans qu'ils aient à échanger les détails de leurs topologies respectives. Comparativement, les protocoles de routage de l'Internet ont également été définis avec une notion de hiérarchie, mais seulement sur 2 niveaux. Que ce soit avec les aires dans OSPF ou les régions dans IS-IS, elles ne dépassent pas le niveau 2 et contraignent leur usage en forçant systématiquement le passage par l'aire ou la région centrale. Il n'est alors possible de hiérarchiser un réseau qu'en fédérant des sous réseaux autour d'un réseau central. BGP lui-même offre deux niveaux de hiérarchie selon que l'on utilise la variante interne iBGP ou externe eBGP. Cependant, contrairement au protocole à ingénierie de trafic comme ISI-IS-TE ou OSPF-TE, BGP ne propage pas d'information relative au taux d'occupation des liens.

L'approche proposée comme étude prospective est de s'appuyer sur les travaux d'agrégation pour hiérarchiser IS-IS comme dans PNNI et s'affranchir de BGP pour améliorer le calcul des chemins d'AS et de placement de tunnels par un serveur PCE.

2.1.3. Méta Routeur

Le principe à étudier (voir figure 28) repose sur la modélisation d'un réseau en un seul routeur. Une agrégation de type maillé « Full-Mesh » est réalisée à partir des données de topologie et de QoS collectées par le protocole IS-IS-TE. Un modèle en étoile « Star » est ensuite construit à partir de cette première modélisation en suivant la méthodologie énoncée précédemment. Soit le méta routeur MR, résultat de cette agrégation. Le cœur du méta routeur MR est représenté par le nucleus du modèle étoile et la capacité de routage du méta routeur MR correspond à la capacité du nucleus en terme d'admission d'appel tel que décrit dans [138]. Les interfaces externes des nœuds de bordure, i.e. les routeurs de périphérie du réseau, représentent alors les interfaces de MR.

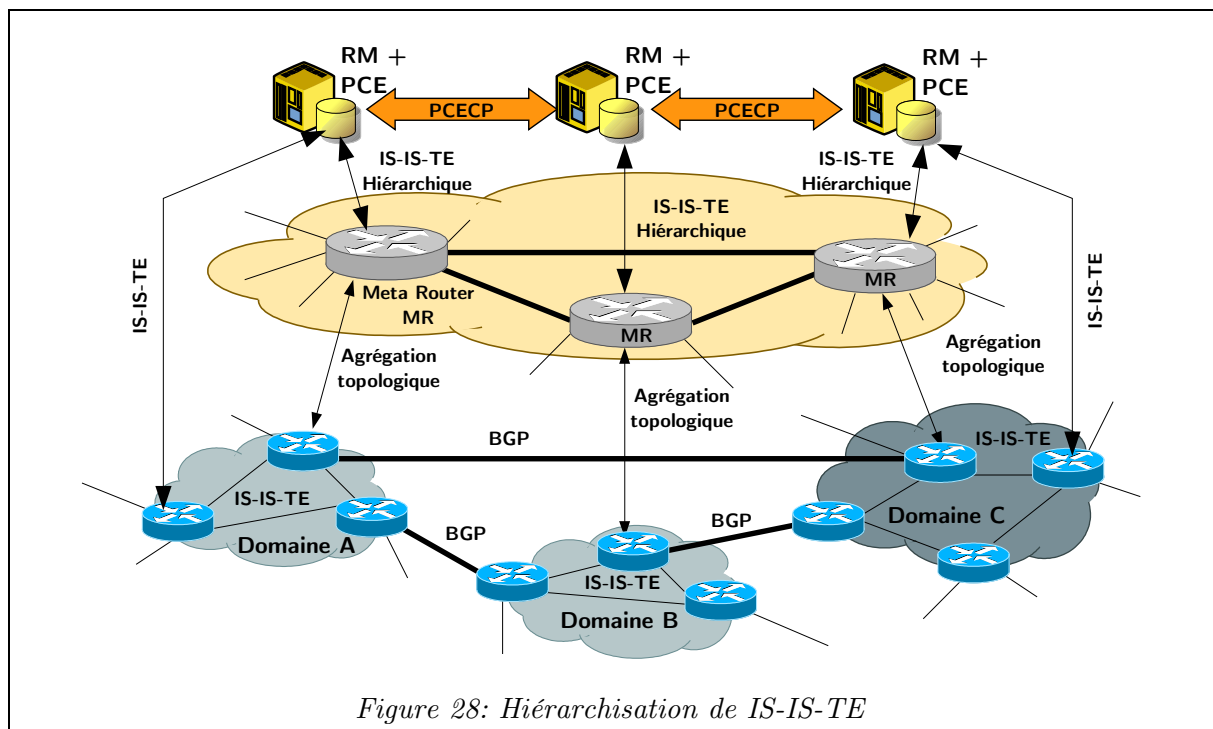
L'ensemble des métas routeurs MR modélisant chacun un réseau constitue un nouveau réseau de hiérarchie supérieure. Ils peuvent ainsi être configurés de sorte à échanger des informations de routage via IS-IS-TE. Dans ce cas, ce sont les informations de QoS calculées sur le modèle de réseau agrégé qui sont véhiculées par le protocole. Chaque méta routeur se comporte alors comme un routeur standard et reconstitue la topologie de niveau supérieur dans sa base de donnée topologique TED (Traffic Engineering Database).

Une fois le protocole de routage à ingénierie de trafic hiérarchisé, il peut être appliqué de façon similaire au serveur PCE qui alimentera sa base de donnée TED via un échange IS-IS-TE avec le méta routeur MR. Une solution optimale consiste à exécuter le processus d'agrégation sur le même équipement que le serveur PCE. Ainsi, lors d'une demande de tunnel inter-domaine, l'algorithme du BRPC utilisera la TED hiérarchique afin de déterminer le meilleur chemin d'AS possible en fonction des disponibilités et des critères de QoS

demandés.

L'agrégation de topologie apporte deux propriétés fondamentales qui rendent le processus de hiérarchisation possible :

- La réduction de la volumétrie des informations de topologie échangées,
 - La réduction de la fréquence des changements de topologie et de routage,
- Les travaux à entreprendre doivent :
- Vérifier les propriétés d'agrégations des réseaux et la modélisation en méta routeur,
 - Étudier et proposer des améliorations d'IS-IS-TE à l'IETF pour échanger des informations de topologies agrégées,
 - Simuler le fonctionnement du protocole IS-IS-TE sur un réseau de méta routeurs pour vérifier la stabilité et mesurer le temps de convergence du protocole dans le calcul de chemin d'AS à QoS garantie,
 - Comparer les propriétés du protocole IS-IS-TE hiérarchisé avec les autres solutions de calcul de chemin d'AS.
 - Étudier l'application du routage hiérarchique à ingénierie de trafic au PCE pour le placement de tunnels inter-domaines.



2.2. Évolution des algorithmes de CAC pour les réseaux domestiques

Le réseau domestique est un cas particulier de réseau très difficile à appréhender pour un opérateur. En effet, un opérateur se doit de garantir un excellent niveau de qualité pour les services qu'il vend tout en n'ayant pas ou peu de maîtrise sur le réseau domestique. D'une part parce qu'il n'en est pas le propriétaire ni le gestionnaire : le client est responsable de son réseau domestique tout en n'ayant pas ou très peu de connaissances en matière de télécommunication. D'autre part parce que l'environnement du réseau domestique est

fortement perturbé et très variable.

Les opérateurs ont résolu partiellement le premier point en louant ou vendant les équipements spécifiques à même de délivrer correctement les services qu'ils vendent. Ainsi au fil du temps, nombre de « Box » intégrant à la fois la partie modem ADSL, routeur, point d'accès WiFi, adaptation RTC/VoiP, décodeur pour accéder à la TV, ... sont apparues sur le marché. En maîtrisant les équipements du réseau domestique, les opérateurs sont sûrs de pouvoir délivrer et garantir le niveau de QoS exigé par leur service. En parallèle, les acteurs du secteur (opérateurs et industriels) s'impliquent de plus en plus dans la normalisation afin de définir de nouveaux standards offrant une interopérabilité des équipements de la sphère domestique. Il est clair que la réduction des coûts passera par une généralisation et une standardisation des « Box » afin de tirer profit de la production en volume. Le futur du réseau domestique va également dans le sens de plus en plus d'équipements connectés comme le réfrigérateur, le carillon avec Webcam intégrée, les alarmes, ... ce que l'on appelle couramment la domotique.

Le deuxième défi auquel les opérateurs doivent faire face concerne la prise en compte de l'environnement très perturbé des réseaux domestiques. Cela touche essentiellement les technologies de transferts WiFi et PLT. La première est sujette aux perturbations radio, tel que le micro-onde ou le voisinage des autres points d'accès (e.g. « Box » disposée de part et d'autre d'un mur séparant deux appartements voisins dans un même immeuble). La deuxième est elle dépendante des équipements directement branchés sur le secteur (e.g. lampe halogène, machine à lavé) qui perturbent la transmission lorsqu'ils sont allumés. Dans les deux cas de figure, le lien de transmission n'est pas stable dans le temps et est sujet à de fortes variations dont il faut pouvoir tenir compte dans la fonction d'admission d'appel.

La période d'étude qui s'ouvre devra donc :

- Étudier les interfaces disponibles sur les équipements du réseau domestique, externes à la passerelle domestique, pour acquérir en temps réel ces changements de débit,
- Étudier de nouveaux algorithmes d'admission d'appel tenant compte de la variabilité des liens de transferts.

2.3. La préemption et les interactions de service

L'interaction de service avec le réseau se borne dans la majorité des cas à notifier l'utilisateur que le service n'est pas disponible en cas de rejet de la communication par la fonction d'admission d'appel. Cependant, dans le cas où une congestion se produit (e.g. lorsque le lien de raccordement se dégrade), il faut pouvoir notifier cette dégradation afin que l'application puisse prendre les mesures appropriées, telle que la réduction du débit par ajustement du codec, l'arrêt de flux de rehaussement, le repli sur un mode dégradé ou plus simplement la notification du problème vers l'utilisateur. Inversement, une application surveillant le niveau de la QoS peut souhaiter notifier le réseau d'une dégradation de la transmission afin que ce dernier corrige les défauts. Finalement, plutôt que de rejeter simplement un nouvel appel, il peut être intéressant de préempter une session en cours afin de libérer des ressources pour un service plus prioritaire e.g. un appel d'urgence.

Les travaux futurs devront s'orienter :

- Vers l'étude et la définition d'une API globale permettant d'une part pour une application de spécifier son niveau de QoS et de notifier le réseau d'une dégradation de la QoS et d'autre part pour le réseau de notifier et/ou préempter les applications si les ressources du réseau viennent à se raréfier,
- Étudier l'interaction entre service et réseau, et plus particulièrement comment

respecter un contrat de service et comment coupler la qualité perçue (QoE) et la qualité de service (QoS).

2.4. Le partage multimédia

Avec la montée en débit des accès Internet auprès du grand public grâce à la fibre optique, de nouvelles perspectives d'applications à forte valeur ajoutée seront offertes. Parmi celles-ci, le partage de contenus multimédia auto produits tient la place la plus importante. Il est ainsi tout à fait concevable de pouvoir, à l'horizon 2-3 ans, échanger en temps réel des contenus audio-visuels à haute qualité entre réseaux domestiques ou en situation de nomadisme.

Cependant, le passage du scénario au service va demander un travail de recherche notamment sur la commande du réseau et plus particulièrement sur la mise en place automatique de la QoS. En effet, ce type de service va nécessiter non seulement de la bande passante – de l'ordre de 10 à 30 Mbit/s selon le type et l'encodage du contenu multimédia – mais surtout, cette bande passante devra être configurée dans le sens montant du client vers le réseau. Ceci est un nouveau défi pour les opérateurs qui ont massivement construit leur réseau sur le modèle dit surdimensionné basé sur un débit montant bridé au limite de la technologie ADSL i.e. au maximum 800 Kbit/s et ne sont pas préparés pour des débits 10 à 100 fois supérieurs. De plus, il n'est pas nécessaire de construire un réseau capable d'absorber de tels débits pour tous les utilisateurs si le taux d'utilisation simultané est faible. A contrario, sous dimensionner son réseau ne permettra pas à un opérateur d'appréhender une adhésion massive à ces nouveaux services.

Les futures études devront donc porter sur :

- La gestion dynamique de la QoS et la configuration dynamique des ressources, qui devront être à la fois du type adressage, constitution de réseaux sociaux privés ou communauté et, bien sûr, la configuration des ressources de transfert,
- La définition de composantes inter réseaux, en particulier domotiques, avec des contraintes de sécurités et de QoS,
- La valorisation des réseaux d'infrastructures des opérateurs (type IMS) pour ne pas qu'ils soient cantonnés au rôle de simples fournisseurs de connectivité.

V

Bibliographie

- [1] Y.1541, *Network Performances Objectives for IP-based Services*, ITU-T, January, 2005
- [2] I.356, *B-ISDN ATM layer cell transfer performance*, ITU-T, May, 1996
- [3] J. Babiarz, K. Chan and F. Baker, *Configuration Guidelines for DiffServ Service Classes*, RFC 4594, June, 2006
- [4] D. Verma, *Service Level Agreements on IP Networks*, Proceedings of the IEEE, Volume 92, (9), September, 2004
- [5] D. Goderis et al., *Service Level Specification Semantics, Parameters and negotiation requirements*, draft-tequila-sls-02.txt, February, 2002
- [6] M. Menaï and G. Fromentoux, *Outil de modélisation, de conception et d'optimisation de l'urbanisme de l'architecture de commande d'un réseau de télécom.*, INPI 05 05557, , 2005
- [7] T. Vu Duong, G. Fromentoux and J.L. Le Roux, *A Global Framework for Architecture Analysis in Telecommunications*, In Proc. Globecom 2003, San Francisco, December, 2003
- [8] Tele Management Forum, *enhanced Telecom Operations Map*, <http://www.tmforum.org>
- [9] M.3400, *TMN Management Function*, ITU-T, July, 2004
- [10] ETSI TISPAN, *Quality of Service (QoS) Framework and Requirements*, ETSI TS 185 001, November, 2005
- [11] P.800.1, *Mean Opinion Score (MOS) terminology*, ITU-T, July, 2006
- [12] D. Harrington, R. Presuhn, and B. Wijnen, *An architecture for describing simple network management protocol (snmp) management frameworks*, RFC 3411, December, 2002
- [13] R. Enns, *Netconf Configuration Protocol*, RFC 4741, December, 2006
- [14] P. Calhoun, J. Loughney, E. Guttman, G. Zorn and J. Arkko, *Diameter Base Protocol*, RFC3588, September, 2003

- [15] D. Durham, J. Boyle, R. Cohen, S. Herzog, R. Rajan, and A. Sastry, *The COPS (Common Open Policy Service) Protocol*, RFC 2748, January, 2000
- [16] I.610, *B-ISDN Operation and Maintenance principles and functions*, ITU-T, February, 1999
- [17] J. Quittek, T. Zseby, B. Claise and S. Zander, *Requirements for IP Flow Information Export (IPFIX)*, RFC 3917, October, 2004
- [18] Martin de Prycker, *Asynchronous Transfer Mode: Solution for Broadband ISDN*, Prentice Hall, 1995
- [19] I.371, *Traffic control and congestion control in B-ISDN*, ITU-T, May, 1996
- [20] ATM Forum, *Traffic Management 4.1*, af-tm-0121.000, March, 1999
- [21] ATM Forum, *ATM User Network Interface (UNI) Signalling Specification version 4.1*, af-sig.0061.002, April, 2002
- [22] ATM Forum, *Private Network-Network Interface Specification v.1.1*, af-pnni-0055.002, April, 2002
- [23] ATM Forum, *PNNI Augmented Routing (PAR) Version 1.0*, af-ra-0104.000, January, 1999
- [24] R. Braden, D. Clark and S. Shenker, *Integrated services in the internet architecture: an overview*, RFC 1633, June, 1994
- [25] S. Shenker, C. Partridge and R. Guerin, *Specification of Guaranteed Quality of Service*, RFC 2212, September, 1997
- [26] J. Wroclawski, *Specification of the Controlled-Load Network Element Service*, RFC 2211, September, 1997
- [27] R. Braden, L. Zhang, S. Berson, S. Herzog, and S. Jamin, *Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification*, RFC 2205 (Updated by RFCs 2750, 3936, 4495), September, 1997
- [28] S. Blake, D. Black, M. Carlson, E. Davies, Z. Wang, and W. Weiss, *An Architecture for Differentiated Service*, RFC 2475 (updated by RFC 3260), December, 1998
- [29] K. Nichols, S. Blake, F. Baker, and D. Black, *Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers*, RFC 2474 (Updated by RFCs 3168, 3260), December, 1998
- [30] B. Davie et al., *An Expedited Forwarding PHB (Per-Hop Behavior)*, RFC 3246, March, 2002
- [31] J. Heinanen, F. Baker, W. Weiss, J. Wroclawski, *Assured Forwarding PHB Group*, RFC 2597 (Updated by RFC 3260), June, 1999
- [32] K. Nichols, V. Jacobson and L. Zhang, *A Two-bit Differentiated Services Architecture for the Internet*, RFC 2638, July, 1999
- [33] P. Chimento et al., *QBone Signaling Team Final Report*, QBone, July, 2002
- [34] O. Schelen, *Quality of Service Agents in the Internet*, Ph.D Thesis from Lulea University of Technology, 1998
- [35] S. Herzog, J. Boyle, R. Cohen, D. Durham, R. Rajan and A. Sastry, *COPS usage for RSVP*, RFC 2749, January, 2000
- [36] R. Yavatkar, D. Pendarakis and R. Guerin, *A Framework for Policy-based Admission Control*, RFC 2753, January, 2000
- [37] R. Sahita, S. Hahn, K. Chan and K. McCloghrie, *Framework Policy Information Base*, RFC 3318, March, 2003

-
- [38] K. Chan et al., *COPS Usage for Policy Provisioning (COPS-PR)*, RFC 3084, March, 2001
 - [39] K. Chan, R. Sahita, S. Hahn and K. McCloghrie, *Differentiated Services Quality of Service Policy Information Base*, RFC 3317, March, 2003
 - [40] E. Rosen, A. Viswanathan and R. Callon, *Multiprotocol Label Switching Architecture*, RFC 3031, January, 2001
 - [41] V. Fuller and T. Li, *Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan*, RFC 4632 Obsoletes RFC 1519, August, 2006
 - [42] E. Rosen et al., *MPLS Label Stack Encoding*, RFC 3032 (Updated by RFCs 3443, 4182), January, 2001
 - [43] L. Anderson, P. Doolan, N. Feldman, A. Fredette, and B. Thomas, *LDP Specification*, RFC 3036, January, 2001
 - [44] B. Davie et al., *MPLS using LDP and ATM VC Switching*, RFC 3035, January, 2001
 - [45] F. Le Faucheur et al., *Multi-Protocol Label Switching (MPLS) Support for Differentiated Services*, RFC 3270, May, 2002
 - [46] Y. Rekhter and E. Rosen, *Carrying Label Information in BGP-4*, RFC 3107, May, 2001
 - [47] D. Awduche, L. Berger, D. Gan, T. Li, V. Srinivasan, and G. Swallow, *RSVP-TE: Extension to RSVP for LSP tunnels*, RFC 3209 (Updated by RFCs 3936, 4420, 4874), December, 2001
 - [48] B. Jamoussi et al., *Constraint-Based LSP Setup using LDP*, RFC 3212 (Updated by 3468), January, 2002
 - [49] D. Katz, K. Kompella, and D. Yeung, *Traffic Engineering (TE) Extensions to OSPF version 2*, RFC 3630 (Updated by RFC 4203), September, 2003
 - [50] H. Smit and T. Li, *Intermediate System to Intermediate System (IS-IS) Extension for Traffic Engineering (TE)*, RFC 3784 (Updated by RFC 4205), June, 2004
 - [51] D. Awduche, J. Malchom, J. Agogbua, M. O'Dell and J. McManus, *Requirements for Traffic Engineering Over MPLS*, RFC 2702, September, 1999
 - [52] F. Le Faucheur, *Protocol Extensions for Support Diffserv-aware MPLS Traffic Engineering*, RFC 4124, June, 2005
 - [53] F. Le Faucheur and W. Lai, *Maximum Allocation Bandwidth Constraints Model for DiffServ-aware MPLS Traffic Engineering*, RFC 4125, June, 2005
 - [54] F. Le Faucheur, *Russian Dolls Bandwidth Constraints Model for DiffServ-aware MPLS Traffic Engineering*, RFC 4127, June, 2005
 - [55] J. Ash and J.L. Le Roux, *Path Computation Element (PCE) Communication Protocol Generic Requirements*, RFC 4657, September, 2006
 - [56] J.P.V.A. Farrel and J. Ash, *A Path Computation Element (PCE)-Based Architecture*, RFC 4655, August, 2006
 - [57] Y. Rekhter, T. LI and S. Hares, *A Border Gateway Protocol 4 (BGP-4)*, RFC 4271, January, 2006
 - [58] IST/FP5 Project, *AQUILA: Adaptive Resource Control for QoS Using an IP-based Layered Architecture*, <http://www-st.inf.tu-dresden.de/aquila/>
 - [59] T. Engel et al., *AQUILA: adaptive resource control for QoS using an IP-based layered architecture*, IEEE Communication Magazine, Volume 41, (1), January,

2003

- [60] IST/FP5 Project, *MESCAL: Management of End-to-end Quality of Service Across the Internet at Large*, <http://www.ist-mescal.org>
- [61] P. Levis, M. Boucadair, P. Morand, G. Pavlou, P. Trimintzios and J. Spencer, *A new perspective for a global QoS-based internet*, Journal of Communications Software and Systems, Volume 1, (1), January, 2005
- [62] D. Griffin et al., *Interdomain routing through QoS-class planes*, IEEE Communication Magazine, Volume 45, (2), February, 2007
- [63] J. Scudder and C. Appanna, *Multisession BGP*, draft-ietf-idr-bgp-multisession-03.txt, January, 2007
- [64] R. Hancock, G. Karagiannis, J. Loughney, and S. Van den Bosch, *Next Steps in Signaling (NSIS): Framework*, RFC 4080, June, 2005
- [65] H. Schulzrinne and R. Hancock, *GIST: General Internet Signalling Transport*, draft-ietf-nsis-ntlp-16, July,
- [66] J. Manner, G. Karagiannis, and A. McDonald, *NSLP for Quality-of-Service Signaling*, draft-ietf-nsis-qos-nslp-16.txt, February, 2008
- [67] A. Bader, L. Westberg, G. Karagiannis, C. Kappler and T. Phelan, *RMD-QOSM - The Resource Management in Diffserv QOS Model*, draft-ietf-nsis-rmd-13.txt, July,
- [68] S. Oueslati and J. Roberts, *A new direction for quality of service: Flow aware networking*, In Proc. NGI'05, Rome, April, 2005
- [69] M. Menth and F. Lehrieder, *Performance Evaluation of PCN-Based Admission Control*, In Proc. 16th International Workshop on Quality of Service, 2008. IWQoS'08, Enschede, June, 2008
- [70] P. Newman, *Backward Explicit Congestion Notification for ATM local networks*, In Proc. Globecom'93, Houston, December, 1993
- [71] J. Song, M. Chang, S. Lee and J. Joung, *Overview of ITU-T NGN QoS Control*, IEEE Communication Magazine, Volume 45, (9), September, 2007
- [72] ETSI TISPAN, *IP Multimedia Subsystem (IMS); Functional architecture*, ETSI ES 282 007 V2.1.1, September, 2008
- [73] ETSI TISPAN, *Resource and Admission Control Sub-system (RACS); Functional Architecture*, ETSI ES 282 003 V2.0.0, May, 2008
- [74] Y.2012, *Functional requirements and architecture of the NGN*, ITU-T, August, 2006
- [75] Y.2111, *Resource and Admission Control Functions in NGN*, ITU-T, August, 2006
- [76] P. Levis and M. Boucadair, *Considerations of Provider-to-Provider Agreements for Internet-Scale Quality of Service (QoS)*, RFC5160, March, 2008
- [77] P. Boyer, *A congestion control for the ATM*, In Proc. 7th ITC Specialists Seminar, Morristown, USA, October, 1990
- [78] P. Boyer and F. Guillemin and M. Serval and J.P. Coudreuse, *Spacing cells protects and enhances utilization of ATM network links*, IEEE Networks, Volume 6, (5), September, 1992
- [79] J. Roberts et al., *Performance evaluation and design of multiservice networks*, Cost 224, Final Report, 1992
- [80] J. Roberts, *Jitter due to an ATM multiplex. Application to peak rate policing*, Cost 224, Chapter 1, pages 35 TD (89), 1989
- [81] J. Turner, *New directions in communications (or which way in the information*

- age ?), IEEE Communications Magazine, Volume 24, (10), October, 1986
- [82] J. Roberts, *Virtual Spacing for flexible traffic control*, Int'l J. Digital and Analog Commun., Volume 7, (), , 1994
 - [83] P. Boyer and M. Servel and F. Guillemin, *The Spacer-Controller: an Efficient UPC/NPC for ATM Networks*, In Proc. of ISS'92, Yokohama, Japan, October, 1992
 - [84] M. Servel, P. Gonet and J. François, *Circuit pour mémoriser des états de disponibilités de ressources logiques*, Brevet INPI n° 87 09068, , 1987
 - [85] E. Wallmeier and T. Worster, *The spacing policer, an algorithm for efficient peak bit rate control in ATM networks*, In , Yokohama, October, 1992
 - [86] J. Roberts and P. Boyer and M. Servel, *A real time sorter with application to ATM traffic control*, In Proc. of ISS'95, Berlin, April, 1995
 - [87] J. Chao, *A novel architecture for queue management in the ATM network*, IEEE Journal Selec. Areas in Commun., Volume , (9(7)), September, 1991
 - [88] D. Knuth, *The Art of Computer Programming, Seminumerical Algorithms*, , Chapter 5 : Sorting, pages 139--159, 1969
 - [89] O. Dugeon, *Machine d'admission pour le Réseau d'Infrastructure ATM*, Ph.D Thesis from University of Rennes I, 1997
 - [90] P. Boyer, O. Dugeon et M. Servel, *Systèmes d'allocation d'intervalle de temps et multiplexeurs pourvus d'un de ces systèmes d'allocation d'intervalle de temps*, Brevet INPI n° 93 09645, Novembre, 1993
 - [91] P. Boyer and M. Servel, *A spacer-multiplexer for public UNIs*, In Proc. ISS'95, Berlin, April, 1995
 - [92] O. Bonaventure, *A simulation study of TCP with the proposed GFR service category*, In Proc. High Performance Networks for Multimedia Applications, Dagstuhl, April, 1997
 - [93] W.R.. Stevens, *TCP/IP illustrated Vol. 1*, Addison Wesley Professional Computing Series, 1994
 - [94] M. Allman, V. Paxson and W. Stevens, *TCP Congestion Control*, RFC 2581 Updated by RFC 3390, April, 1999
 - [95] F. Guillemin, J. Boyer, O. Dugeon and C. Mangin, *Analysis of cell spacing on a TCP flow mapped onto an ATM connection*, In Proc. Networking 2000, Paris, May, 2000
 - [96] F. Guillemin, J. Boyer, P. Boyer, O. Dugeon and C. Mangin, *Cell spacing for regulating TCP over ATM*, In Proc. International Teletraffic Congress (ITC'16), , June, 1999
 - [97] Hong-Bin Chiou, Sheng-Der Chin and Zsehong Tsai, *Enhancing the fairness of TCP over Internet using an improved hierarchical packet fair queueing scheme*, In Proc. of IEEE International Conference on Networks, , September, 2000
 - [98] R.Takano, T.Kudoh, Y.Kodama, M.Matsuda, H.Tezuka, and Y.Ishikawa, *Design and Evaluation of Precise Software Pacing Mechanisms for Fast Long-Distance Networks.*, In Proc. 3rd Intl. Workshop on Protocols for Fast Long-Distance Networks (PFLDnet05), Lyon, February, 2005
 - [99] Y.1221, *Traffic control and congestion control in IP based networks*, ITU-T, March, 2002
 - [100] J. Boyer, P. Boyer, O. Dugeon, F. Guillemin, and C. Mangin, *Accelerated signalling*

- for the internet over ATM, European Transactions on Telecommunications (ETT), Volume 10, (2), March/April, 1999
- [101] G. Hjálmtýsson and KK Ramakrishnan, *UNITE - An architecture for lightweight signalling in ATM networks*, In Proc. Infocom'98, Boston, April, 1998
 - [102] I. Cidon et al., *OPENET: an open and efficient control platform for ATM networks*, In Proc. INFOCOM '98, , April, 1998
 - [103] B. Qiyong, K. Shiimoto and J. Turner, *Dynamic flow switching. A new communication service for ATM networks*, In Proc. INFOCOM'98, , April, 1998
 - [104] G. Guillemin, O. Dugeon, J. Boyer and C. Mangin, *Lightweight signaling in ATM networks for high quality transfer of Internet traffic*, Computer Networks, Volume 34, (2), August, 2000
 - [105] F. Guillemin, O. Dugeon, C. Mangin, H. Zhang-Li, M. Basham & N. Dangy-Caye, *Enrichment of MPLS over ATM for guaranteeing quality of service*, In Proc. International Switching Symposium (ISS'2000), Birmingham, May, 2000
 - [106] O. Dugeon, F. Guillemin et C. Mangin, *Contrôle d'admission des flots TCP dans un réseau ATM par signalisation mono cellule*, Brevet INPI n°99 14722, Novembre, 1999
 - [107] G. Cortese et al., *Cadenus: creation and deployment of end-user services in premium IP networks*, IEEE Communication Magazine, Volume 41, (1), January, 2003
 - [108] O. Dugeon and A. Diaconescu, *From SLA to SLS up to QoS control : The CADENUS Framework*, In Proc. World Telecommunication Congress, WTC/ISS'02, Paris, France, September, 2002
 - [109] G. Cortese (Ed.), *Mediation Component Release 1 -Requirements and Architecture*, Cadenus Project, Deliverable D3.1, Deliverable , 2001
 - [110] T. Nguyen, N. Boukhatem and G. Puiolle, G., *COPS-SLS usage for dynamic policy-based QoS management over heterogeneous IP networks*, IEEE Networks, Volume 17, (3), May-June, 2003
 - [111] G. Cortese (Ed.), *Service Configuration and Provisioning Framework: Requirements, Architecture, Design*, Cadenus Project, Deliverable D3.2, 2001
 - [112] Eurescom, *Selected Scenarios and requirements for end-to-end IP QoSmanagement*, Project P1008, Deliverable D2, 2001
 - [113] O. Dugeon (Ed.), *Service Provisioning in Premium IP: Recommendations to Telecom Operators and ISPs*, Cadenus Project, Deliverable D6.2, 2001
 - [114] O. Dugeon and A. Brajeul (Ed.), *Project Recommendations*, Cadenus Project, Deliverable D6.3, 2003
 - [115] G. Ash, A. Brader, C. Kappler, D. Oran, *QoS NSLP QSPEC Template*, draft-ietf-nsis-qspec-20.txt, April, 2008
 - [116] M. van Hartskamp et al., *UPnP QoS Architecture*, UPnP Forum, October, 2006
 - [117] IST/FP6 Project, *EuQoS: End-to-End Quality Service Over Heterogeneous Networks*, <http://www.euqos.eu>
 - [118] T. Braun, M. Diaz, J.E. Gabeiras and T. Staub, *End-to-End Quality of Service Over Heterogeneous Network*, Springer, 2008
 - [119] O. Dugeon et al., *End to End Quality of Service over Heterogeneous Networks*, In Proc. Networking Control, Lannion, France, November, 2005
 - [120] A. Beben, *EQ-BGP: an efficient inter-domain QoS routing protocol*, In Proc.

- Advanced Information Networking and Applications, AINA'2006, , April, 2006
- [121] Olivier Dugeon, Enzo Mingozzi, Giovanni Stea, Luca Bisti, *Provisioning QoS in Inter-domain Traffic Engineering*, Annals Of Telecommunications, Volume 63, (11-12), November, 2008
 - [122] F. Luyuan, N. Bitar, J-L. Le Roux, and J. Miles, *Interprovider IP-MPLS services: Requirements, Implementations and Challenges*, IEEE Communication Magazine, Volume 43, (6), June, 2005
 - [123] A. Farrel, J.P. Vasseur and A. Ayyangar, *A Framework for Inter-Domain Multiprotocol Label Switching Traffic Engineering*, RFC 4726, November, 2006
 - [124] K. Kompella and Y. Rekhter, *Label Switched Paths (LSP) Hierarchy with Generalized Multi-Protocol Label Switching (GMPLS) Traffic-Engineering (TE)*, RFC 4206, October, 2005
 - [125] J.L. Le Roux, J.P. Vasseur, Y. Ikejiri and R. Zhang, *IS-IS Protocol Extensions for Path Computation Element (PCE) Discovery*, RFC 5089, January, 2008
 - [126] J.L. Le Roux, J.P. Vasseur, Y. Ikejiri and R. Zhang, *OSPF Protocol Extensions for Path Computation Element (PCE) Discovery*, RFC 5088, January, 2008
 - [127] J.P. Vasseur, R. Zhang, N. Bitar and J.L. Le Roux, *A Backward Recursive PCE-based Computation (BRPC) Procedure To Compute Shortest Constrained Inter-domain TE LSP*, draft-ietf-pce-brpc-09.txt, April, 2008
 - [128] K. Argyraki and D. Cheriton, *Loose source routing as a mechanism for traffic policies*, In Proc. ACM SIGCOMM workshop on Future Directions in Network Architecture, Portland, August, 2004
 - [129] RIPE, *Weekly Routing Table Report*, <http://www.ripe.net>
 - [130] X. Masip-Bruin et al., *The EuQoS system: a solution for QoS routing in heterogeneous networks*, IEEE Communication Magazine, Volume 45, (2), February, 2007
 - [131] H.C Lin, S.C Lai and P.W Chen, *Automatic topology discovery of IP networks*, In Proc. IEICE Transactions, , January, 2000
 - [132] Cisco, *Cisco Discovery Protocol*, <http://www.cisco.com>
 - [133] Hewlett-Packard, *HP OpenView*, <http://www.openview.hp.com>
 - [134] Paul Congdon and Bill Lane, *Station and Media Access Control Connectivity Discovery*, IEEE 802.1AB, April, 2005
 - [135] W. Htira, O. Dugeon and M. Diaz, *STAMP: Towards A Scalable Topology Announcement and Management Protocol*, In Proc. of 22nd IEEE International Conference on Advanced Information Networking and Applications (AINA2008), Okinawa, March, 2008
 - [136] L. Guo and I. Matta, *On State Aggregation for Scalable QoS Routing*, In Proc. ATM Workshop'98, , May, 1998
 - [137] W. Htira, O. Dugeon and M. Diaz, *A New Approximation Model for Guaranteed QoS Model Information*, In Proc. IEEE Symposium on Computers and Communications ISCC'07, Portugal, July, 2007
 - [138] W. Htira, O. Dugeon and M. Diaz, *An Aggregated Delay/Bandwidth Star Scheme For Admission Control*, In Proc. of 18th Annual IEEE International Symposium on Personal, Indoor and Radio Communications – (PIMRC'07), Athens, September, 2007

- [139] W. Htira, O. Dugeon and M. Diaz, *A Mesh Aggregation Scheme For Call Admission Control*, In Proc. of Australasian Telecommunication Networks and Applications Conference (ATNAC'07), New-Zealand, December, 2007
- [140] W. Htira, O. Dugeon and M. Diaz, *A Novel Bandwidth Broker Architecture Based on Topology Aggregation in Delay/Bandwidth Sensitive Networks*, In Proc. IFIP, Networking 2008, Singapour, May, 2008
- [141] W. Htira, *Découverte et Agrégation de Topologies de Réseaux. Application au Contrôle d'Admission*, Ph.D Thesis from University of Toulouse, 2008
- [142] HGI, *Home Gateway Requirements: Residential Profile*, Home Gateway Initiative, April, 2008
- [143] O. Dugeon, *CAC requirements for the Home Gateway*, HGI contribution, May, 2007
- [144] O. Dugeon, *Procédé de traitement d'une requête, passerelle d'accès et système de contrôle d'admission à un service*, Brevet INPI n° 08 54383, Juin, 2008