



HAL
open science

Architecture à qualité de service pour systèmes satellites DVB-S/RCS dans un contexte NGN

Olivier Alphand

► **To cite this version:**

Olivier Alphand. Architecture à qualité de service pour systèmes satellites DVB-S/RCS dans un contexte NGN. Réseaux et télécommunications [cs.NI]. Institut National Polytechnique de Toulouse - INPT, 2005. Français. NNT: . tel-00011402

HAL Id: tel-00011402

<https://theses.hal.science/tel-00011402>

Submitted on 17 Jan 2006

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.

THESE

Préparée au

Laboratoire d'Analyse et d'Architecture des Systèmes du CNRS

En vue de l'obtention du titre de

Docteur de l'Institut National Polytechnique de Toulouse
Spécialité : Informatique et télécommunications

**ARCHITECTURE A QUALITE DE SERVICE
POUR SYSTEMES SATELLITES DVB-S/RCS
DANS UN CONTEXTE NGN**

Par
Olivier ALPHAND

Soutenue le 07/12/2005 devant le jury :

Président du jury
Directeur de thèse
Rapporteur
Rapporteur
Examineur, Co-Encadrant
Examineur, Co-Encadrant
Membre Invité

Christian FRABOUL
Michel DIAZ
Guy PUJOLLE
François LEPAGE
Thierry GAYRAUD
Pascal BERTHOU
Stéphane COMBES

Remerciements

Les travaux présentés dans ce mémoire ont été réalisés dans le cadre d'une collaboration entre le Laboratoire d'Analyse et d'Architecture des Systèmes du CNRS et Alcatel Alenia Space à travers successivement le projet Européen SATIP6 (Nov. 2002-Avril 2004) et une convention entre AAS et le LAAS (Mai 2004 – Déc. 2005).

La thèse a été effectuée au sein du groupe OLC (Outils Logiciels pour la Communication) dans le Laboratoire d'Analyse et d'Architecture des Systèmes du CNRS sous la direction successive de Mr Jean-Claude Laprie et Mr Malik Ghallab.

Je commencerai par remercier Mr Christian Fraboul qui m'a fait l'honneur d'assumer la fonction de président de jury de ma thèse. Je suis également reconnaissant et honoré que les professeurs Guy Pujolle et François Lepage aient accepté de rapporter ma thèse et les remercie pour leurs remarques pertinentes sur mon travail.

Je remercie Michel Diaz, mon directeur de thèse, pour m'avoir fait profiter de son expérience et de ses critiques avisées qui m'ont donné un éclairage complémentaire sur mes travaux.

J'exprime une profonde reconnaissance à Stéphane Combes, mon interlocuteur privilégié à Alcatel Alenia Space, pour m'avoir guidé dans ce territoire inconnu qu'était pour moi les systèmes satellites ainsi que pour sa disponibilité, son ouverture, ses multiples conseils et sa relecture de ce mémoire.

Je tiens à remercier Thierry Gayraud de m'avoir offert l'opportunité de découvrir le monde de la recherche à travers mon stage de DEA et de m'avoir encadré tout au long de cette thèse. J'ai notamment apprécié son recul sur mon travail, la confiance qu'il m'a témoignée tout au long de ces trois ans, le contexte de travail de qualité qu'il m'a offert ainsi que l'autonomie qu'il m'a laissée pour explorer de nombreuses pistes de travail.

Je souhaite aussi vivement remercier Pascal Berthou pour son encadrement, son soutien, son « support technique » et ses nombreux encouragements. Sa vision pragmatique, son écoute, les différents recadrages qu'il a effectué pour m'éviter de me disperser et nos longues discussions animées ont largement contribué à faire de ces trois ans une aventure extrêmement enrichissante.

Je remercie également l'ensemble des acteurs du projet SATIP6 et l'équipe d'Alcatel avec qui j'ai réellement apprécié travailler.

Merci également aux membres du groupe OLC, permanents, doctorants et stagiaires, pour leur accueil chaleureux et leur bonne humeur : les deux Nicolas, Fred, Stéphane, Yann, Philippe, Guillaume, Florin, Emir, Christophe, André, Rehan, Jarl, Lode, et tous les autres. Je salue également les doctorants et autres membres du laboratoire Tésa et du département Mathématiques et Informatiques de l'ENSICA. Je remercie aussi les membres des différents services techniques et administratifs du LAAS-CNRS, qui m'ont permis de travailler dans d'excellentes conditions.

Enfin merci encore à tous ceux qui m'ont accompagné et soutenu au cours de ces trois dernières années : ma famille, mes nombreux amis de Toulouse, de Strasbourg et d'ailleurs et Marie que je ne remercierai jamais assez pour son soutien sans faille et sa patience admirable.

Liste des acronymes

3GPP	3rd Generation Partnership Project
AAL5	ATM Adaptation Layer 5
AC	Admission Control
ACQ	ACQuisition burst
ACSS	Access Control SubSystem
ADSL	Asymmetric Digital Subscriber Line
AF	Assured Forwarding
API	Application Programming Interface
AR	Address Resolution
ARC	Access Resource Control
ARP	Address Resolution Protocol
ASP	Application SP
ATM	Asynchronous Transfer Mode
ATSC	Advanced Television Systems Committee
BAS	Broadband Access Server
BAT	Bouquet Association Table
BE	Best Effort
BER	Bit Error Rate
BoD	Bandwidth on Demand
BSM	Broadband Satellite Multimedia
C2P	Connection Control Protocol
CAC	Connection Admission Control
CAT	Conditional Access Table
CBR	Constant Bit Rate
CBR	Constant Bit Rate
CMT	Correction Map Table
COPS	Common Open Policy Service
CPE	Customer Premises Equipment
CRA	Continuous Rate Assignment
CSC	Common Signalling Channel
DAMA	Demand Assignment Multiple Access
DHCP	Dynamic Host Configuration Protocol
DIPCAST	Dvb pour l'IP multiCAST
DSCP	DiffServ Code Point
DSL	Digital Subscriber Line
DSM-CC	Digital Storage Media Command and Control
DULM	Data Unit Labelling Method
DVB	Digital Video Broadcasting
DVB	Digital Video Broadcast
DVB-C	DVB for Cable
DVB-RCS	DVB Return Channel for Satellite
DVB-S	DVB for Satellite
DVB-SI	DVB System Information
DVB-T	DVB for Terrestrial
EDF	Earliest Deadline First
EF	Expedited Forwarding
ER	Edge Router
ES	Elementary Stream
ESA	European Space Agency
ETSI	European Telecommunications Standards Institute

FAI	Fournisseur d'Accès Internet
FCA	Free Capacity Assignment
FCT	Frame Composition Table
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FEC	Forward Error Correcting
FLS	Forward Link Signalling
FLSS	Forward Link SubSystem
FMKE	Flat Multicast Key Exchange
FTP	File Transfer Protocol
GEO	Geostationary Earth Orbit
GMSS	GW Management SubSystem
GPRS	General Packet Radio System
GSM	Global System for Mobile
GW	Gateway
HDTV	High Definition TeleVision
HTTP	Hyper Text Transfer Protocol
IBIS	Integrated Broadcast Interactive System
IE	Information Element
IETF	Internet Engineering Task Force
IGMP	Internet Group Management Protocol
IMS	IP Multimedia SubSystem
IN	Interactive Network
INAP	Interactive Network Access Provider
INT	IP/MAC Notification Table
IP	Internet Protocol
IPD	IP dédié
IPsec	IP security protocol
IRD	Integrated Receiver Decoder
ISDN	Integrated Services Digital Network
ISP	Internet Service Provider
IST	Information Science and Technology
ITSP	Internet Telephony Service Provider
ITU	International Telecommunications Union
JT	Jitter Tolerant
LAN	Local Area Network
LEO	Low Earth Orbit
MAC	Medium Access Control
MCDDD	Multi-Carrier Demultiplexer Demodulator Decoder
MF-TDMA	Multi-Frequency Time Division Multiple Access
MMCS	MultiMedia Call Server
MMT	Multicast Map Table
MNMC	Mission and Network Management Centre
MOS	Mean Opinion Score
MPE	Multi Protocol Encapsulation
MPEG	Motion Picture Expert Group
MPLS	Multi-Protocol Label Switching
MSL	Minimum Scheduling Latency
MSP	Multicast Service Provider
MTU	Maximum Transfer Unit
NAT	Network Address Translation
NCC	Network Control Centre
NGN	Next Generation Network
NIT	Network Information Table

NMC	Network Management Center
NNI	Network-Network Interface
NRT	Non Real Time
OBP	On-Board-Processing
OBPC	On-Board-Processing Controller
OSA	Open Service Access
P2P	Peer-to-Peer
PAT	Program Association Table
PCF	Policy Control Function
PCR	Program Clock Reference
PDB	Per Domain Behavior
PDP	Policy Decision Point
PDU	Protocol Data Unit
PEP	Policy Enforcement Point
PES	Packetized Elementary Stream
PHB	Per-Hop Behavior
PID	Packet Identifier
PIM-SM	Protocol Independent Multicast – Sparse Mode
PMT	Program Map Table
PQ	Priority Queuing
PS	Program Stream
PSI	Program Service Information
PSTN	Public Switched Telephonic Network
PUSI	Payload Unit Start Indicator
PVC	Permanent Virtual Channel
QoS	Quality of Service
RAP	Resource Allocation Protocol
RBDC	Rate-Based Dynamic Allocation
RC-EDF	Rate Controlled Earliest Deadline First
RCS	Return Channel for Satellite
RCST	Return Channel Satellite Terminal
RLSS	Return Link SubSystem
RMT	Rcs Map Table
RSVP	Resource reSerVation Protocol
RT	Real Time
RTC	Réseau Téléphonique Commuté
RTP	Real Time Protocol
RTSP	Real-Time Streaming Protocol
RTT	Round Trip Time
SAC	Satellite Access Control
SACK	Selective ACKnowledgment
SAN	Satellite Access Network
SAR	Segmentation and Reassembly
SARP	Satellite Address Resolution Protocol
SATIP6	Satellite Broadband Multimedia System for IPv6
SCR	Source Clock Reference
SCT	Superframe Composition Table
SDAF	Satellite Dependent Access Function
SDP	Session Description Protocol
SDT	Service Description Date
SE	Satellite Emulator
SI	Service Information
SIAF	Satellite Independent Access Function
SIP	Session Initiation Protocol

SI-SAP	Satellite Independent Service Access Point
SLA	Service Level Agreement
SLS	Service Level Specification
SNDU	SubNetwork Data Unit
SNO	Satellite Operator Network
SOAP	Simple Object Access Protocol
SP	Service Provider
SPT	Satellite Position Table
ST	Satellite Terminal
STC	Source Time Clock
SYNC	SYNChronisation burst
TBTP	Terminal Burst Time Plan
TCB	Traffic Conditioning Block
TCP	Transmission Control Protocol
TCT	Time slot Composition Table
TDM	Time Division Multiplexing
TDMA	Time Division Multiple Access
TIM	Terminal Information Message
TIPHON	Telecommunications and Internet Protocol Harmonization Over Networks
TISPAN	Telecommunications and Internet converged Services and Protocols for Advanced Networking
TS	Transport Stream
UA	User Agent
UAC	User Agent Client
UAS	User Agent Server
UDLR	UniDirectional Link Routing
ULE	Unidirectionnal Lightweight Encapsulation
UMTS	Universal Mobile Telephone Service
UNI	User-Network Interface
URI	Uniform Resource Identifier
VBDC	Volume Based Dynamic Capacity
VBR	Variable Bit Rate
VCI	Virtual Channel Identifier
VoD	Video on Demand
VoIP	Voice over IP
VPI	Virtual Path Identifier
VPN	Virtual Private Network
VR	Variable Rate
VSAT	Very Small Aperture Terminal
XML	eXtensible Markup Language

Liste des figures

Figure 1 : Segmentation des acteurs NGN (source : Arcome)	21
Figure 2 : Architecture générique des NGNs	22
Figure 3 : Session SIP Basique	26
Figure 4 : ITU-T G.1010 - Besoins des applications en terme d'interactivité et de fiabilité	32
Figure 5 : Architecture de contrôle de politique	35
Figure 6 : Modèle hiérarchique de service différencié	36
Figure 7 : Session Q-SIP basique	38
Figure 8 : Etablissement de session avec précondition	40
Figure 9 : Architecture globale de QoS NGN	42
Figure 10 : Débit et portée théoriques des technologies sans fil	43
Figure 11 : Format de la trame 802.11b	44
Figure 12 : Partage de bande passante WiFi entre un hôte lent et un hôte rapide	44
Figure 13 : infrastructure d'accès satellite bidirectionnel basique	49
Figure 14 : Communication inter-ST avec satellite transparent et régénératif	50
Figure 15 : Satellite bidirectionnel multi-spots régénératif	51
Figure 16 : Multiplexage MPEG2-TS	53
Figure 17 : Format d'un paquet MPEG2-TS	54
Figure 18 : Pile protocolaire DVB	55
Figure 19 : Encapsulation d'un datagramme IP selon MPE	56
Figure 20 : Encapsulation ULE	57
Figure 21 : Composition d'une supertrame DVB-RCS	59
Figure 22 : Architecture protocolaire DVB-S/RCS dans le plan de donnée	60
Figure 23 : Pile protocolaire pour la signalisation RCS sur la voie aller	61
Figure 24 : Architecture protocolaire d'un système régénératif	63
Figure 25 : Charge utile régénérative	63
Figure 26 : Segmentation hiérarchique des ressources d'une voie montante	66
Figure 27 : Architecture protocolaire BSM pour service unicast	69
Figure 28 : Architecture générique SATIP6	74
Figure 29 : Entête d'un paquet IP-dédié	76
Figure 30 : Encapsulation d'IP dédié sur AAL5/ATM	76
Figure 31 : Table de résolution d'adresse	77
Figure 32 : Les fonctionnalités d'une Gateway	79
Figure 33 : Structure des SLAs	80
Figure 34 : Interaction entre ST, OBP et NCC dans un scénario maillé	81
Figure 35 : QoS dans l'ER et la GW	82
Figure 36 : Architecture de QoS d'un ST	84
Figure 37 : Structure d'une supertrame composée de n=4 trames	86
Figure 38 : Cycle de Requête/Allocation DAMA	87
Figure 39 : influence d' α sur l'efficacité du DAMA	88
Figure 40 : Saturation de la queue MAC DVB-NRT	89
Figure 41 : Station de contrôle regroupant NCC et NMC	91
Figure 42 : Scénario de démonstration SATIP6	92
Figure 43 : Pile protocolaire SATIP6	94
Figure 44 : Blocs Fonctionnels de l'émulateur satellite	95
Figure 45 : Gestion de QoS par le ST	97
Figure 46 : Interface graphique de l'agent de QoS	97
Figure 47 : Format XML véhiculé par l'agent de QoS	99

Figure 48 : Module de communication (a) et du Traceur de connexions (b) de l'agent de QoS..	100
Figure 49 : Diagramme de classe du serveur de QoS.....	101
Figure 50 : Ssion SIP avec réservation de QoS par le proxy en mode « QoS Enabled ».....	104
Figure 51 : Session SIP avec réservation de QoS par le proxy en mode « QoS Assured »	104
Figure 52 : Format XML communiqué par le proxy SIP	105
Figure 53 : Profil de débit d'audioconférence G711, GSM (Gnomemeeting) et SIREN (Windows Messenger) avec suppression de silence.....	110
Figure 54 : CDF des délais inter-paquet à la source.....	111
Figure 55 : Profil de débit de vidéoconférence sous Gnomemeeting (H261) (a) et Windows Messenger (H263) (b)	112
Figure 56 : Configuration du codeur vidéo H261 de Gnomemeeting	112
Figure 57 : Fonctions empiriques de répartition (a) et de distribution cumulative (b) de taille de paquets	113
Figure 58 : CDF des délais inter-paquet à la source (vidéoConférence).....	113
Figure 59 : Profil de débit des Sessions vidéo à la demande.....	115
Figure 60 : Plate-forme de test	118
Figure 61 : Encapsulation RTP/UDP/IP/AAL5/ATM	120
Figure 62 : Algorithme Token Bucket.....	121
Figure 63 : (a) débit TCP durant la congestion, (b) requêtes de capacité associées.....	124
Figure 64 : Influence du DAMA sur une connexion GSM	125
Figure 65 : influence du DAMA sur un profil MPEG (a, b &c)	126
Figure 66 : a) Débit d'une vidéo à la demande basé sur le codec DIVX (Débit moyen = 400 kbps) b) délai de bout en bout par paquet avec $\alpha=0.5$ et FCA=40 kbps	127
Figure 67 : a) Fonction de densité cumulative des délais de bout en bout avec un FCA = 0 kbps b) Avec un FCA = 40 kbps	127
Figure 68 : Influence de l'agrégation de trafic VoIP homogène sur IPTD	128
Figure 69 : Influence de l'agrégation de trafic VoIP homogène sur IPDV	129
Figure 70 : Changement dynamique de CdS initié par l'agent de QoS	130
Figure 71 : Temps caractéristiques d'établissement d'une session SIP	131

Liste des tableaux

Tableau 1 : Décomposition TIPHON du plan de contrôle.....	24
Tableau 2 : Classes de trafic BSM.....	32
Tableau 3 : Classes de service et priorité inter-classe.....	82
Tableau 4 : Délai, Gigue et Pertes pour les applications audio et vidéo selon la recommandation Y.1540 ITU-T.....	108
Tableau 5 : Caractéristiques des codecs VoIP.....	108
Tableau 6 : Débit IP par type de codeur.....	109
Tableau 7 : Débits associés à différents services audiovisuels.....	111
Tableau 8 : Caractéristiques des vidéos capturés.....	114
Tableau 9: Composition du trafic ADSL par application.....	116
Tableau 10 : Configuration de référence de la couche physique.....	121
Tableau 11 : Configuration de la couche MAC.....	121
Tableau 12 : Configuration de la QoS de la couche IP.....	122
Tableau 13 : Adaptation de la pile TCP au segment satellite.....	123
Tableau 14 : Performance de l'architecture de QoS SATIP6 dans différentes conditions de charge.....	130
Tableau 15 : temps moyen d'aller-retour des messages SIP en fonction de la charge de trafic du ST.....	131

Table des matières

INTRODUCTION GENERALE.....	15
CHAPITRE 1 : LA QOS DANS LES RESEaux DE NOUVELLE GENERATION ...	19
I.1 Introduction	19
I.2 Les réseaux de nouvelle génération	19
I.2.1 Mise en garde : Terminologie	19
I.2.2 L'avènement.....	19
I.2.3 L'architecture NGN.....	21
I.2.4 Conclusion	28
I.3 Architectures de QoS existantes.....	28
I.3.1 Préambule.....	28
I.3.2 Architecture « orientée utilisateur ».....	29
I.3.3 Architecture de QoS « orientée réseau »	30
I.3.4 Une architecture de QoS « orientée session ».....	38
I.4 Architecture de QoS NGN.....	41
I.4.1 Architecture globale de QoS NGN.....	41
I.4.2 QoS dans les réseaux d'accès sans fil: Problématique Multi-réseaux.....	43
I.5 Conclusion.....	45
CHAPITRE 2 : QOS DANS LES RESEaux SATELLITE DE NOUVELLE GENERATION.....	47
II.1 Les satellites, la réponse à la fracture numérique ?.....	47
II.1.1 Les avantages intrinsèques.....	47
II.1.2 Les avancées récentes	48
II.1.3 Les systèmes satellites d'accès bidirectionnels.....	48
II.2 La norme DVB-S et support d'IP	52
II.2.1 La norme DVB-S	52
II.2.2 Méthode d'accès.....	54
II.2.3 Méthode d'encapsulation d'IP sur DVB-S	55
II.3 La norme DVB-RCS	57
II.3.1 Méthode d'accès : MF-TDMA	58
II.3.2 La signalisation dans un système DVB-RCS/S	60
II.3.3 Intelligence embarquée	62
II.3.4 Les connexions.....	66
II.4 Architecture de QoS dans les réseaux d'accès satellite DVB-S/RCS.....	67
II.4.1 Les différents acteurs du réseau satellite	67
II.4.2 Le modèle architecturale BSM basé sur IP.....	68
II.4.3 L'algorithme de DAMA : allocation de bande passante à la demande	69
II.5 Conclusion	70
CHAPITRE 3 : ARCHITECTURE DE QOS POUR SYSTEMES DVB-S/RCS.....	73
III.1 Introduction	73
III.2 Le projet Européen SATIP6.....	73
III.2.1 Les objectifs.....	73
III.2.2 L'architecture SATIP6.....	74
III.3 L'architecture de QoS SATIP6	77
III.3.1 Vue d'ensemble	77
III.3.2 La voie aller : la QoS dans la GW.....	78
III.3.3 La voie retour, la QoS dans le ST.....	83
III.4 La plateforme expérimentale SATIP6.....	92
III.4.1 Le scénario de démonstration	92
III.4.2 La plateforme d'émulation.....	93
III.4.3 L'émulation de la couche physique du segment satellite.....	94

III.4.4 Configuration de la plateforme	95
III.5 QoS dynamique pour les applications non conscientes de la QoS sous-jacente	96
III.5.1 Un agent de sélection pour les applications non adaptées à la QoS	96
III.5.2 Signalisation de QoS par un Proxy SIP.....	102
III.6 Conclusion.....	106
CHAPITRE 4 : EVALUATION DE L'ARCHITECTURE DE QOS DVB-RCS	107
IV.1 Introduction	107
IV.2 Caractérisation des applications multimédias.....	107
IV.2.1 L'enjeu.....	107
IV.2.2 Recensement des applications multimédias.....	108
IV.2.3 Les limites.....	116
IV.2.4 Outil de rejeu et plate-forme de mesures	117
IV.3 Calibration de la plate-forme	119
IV.3.1 Remarques préliminaires.....	119
IV.3.2 La configuration de référence	120
IV.4 Evaluation de QoS	123
IV.4.1 Comportement de TCP.....	123
IV.4.2 L'influence du DAMA sur les sources à débit constant et variable : configuration du paramètre d'anticipation	125
IV.4.3 Evaluation de la QoS offerte par un service de type EF.....	127
IV.4.4 Evaluation globale de l'architecture de QoS.....	129
IV.4.5 Evaluation des performances de l'architecture de QoS basée sur de l'agent de QoS et le proxy SIP.....	130
IV.5 Conclusion.....	131
CONCLUSION GENERALE	133

Introduction Générale

Depuis quelques années, les opérateurs de télécommunication ont progressivement pris conscience de la nécessité de faire évoluer leurs infrastructures de communication. Le succès grandissant d'Internet n'y est pas étranger. Ce succès repose principalement sur la technologie IP qui possède deux atouts majeurs.

Le premier est le rôle fédérateur d'IP qui offre une connectivité de bout en bout indépendamment de la nature des réseaux sous-jacents. Ce premier point constitue une révolution en comparaison avec les solutions monolithiques des architectures de télécommunications traditionnelles.

Le deuxième atout est sa neutralité vis-à-vis des applications puisqu'IP ne suppose rien des services qu'il sera amené à supporter ce qui n'est pas le cas des architectures de télécommunication classiques conditionnées par les services qu'elles supportent.

Fort de ce constat, les modèles des réseaux de nouvelle génération (NGN Next Generation Network) imaginés par les opérateurs de télécommunication sont basés sur IP. Dans ce modèle, ils entendent toutefois surmonter l'un des principaux défauts d'Internet à l'heure actuelle : son incapacité à offrir une Qualité de Service (QoS Quality of Service) de bout en bout. De plus, les opérateurs des réseaux et des télécoms misent également sur un fort développement des technologies d'accès sans fil pour d'une part répondre aux exigences de mobilité des utilisateurs et d'autre part réduire la fracture numérique.

Ce contexte constitue le cadre de référence de nos travaux. A la frontière entre le monde des télécommunications et des réseaux, l'objectif de nos contributions est de faciliter l'intégration des réseaux satellites géostationnaires comme technologie d'accès dans l'infrastructure des NGNs. Ce mémoire entend ainsi proposer une architecture de QoS à la fois adaptée au satellite tout en restant compatible avec l'architecture de QoS des NGNs.

Problématique

Héritant d'une longue tradition de solutions propriétaires fermées, les systèmes satellites ont longtemps été caractérisés par leur forte hétérogénéité d'un constructeur à un autre. De cette hétérogénéité, découlait des problèmes d'interopérabilité et un coût particulièrement important des équipements satellites (embarqués ou au sol).

La tendance actuelle de recourir à des standards ouverts comme DVB-S et plus récemment DVB-RCS conduit progressivement à l'homogénéisation des systèmes satellites. Par ailleurs, les efforts d'interopérabilité menés par l'ensemble de la communauté satellite ont permis de baisser significativement les coûts des équipements. Le succès de la norme DVB-S à travers les services de télévision numérique a permis ainsi de démocratiser les terminaux satellites.

En ce qui concerne le support d'IP, différentes solutions satellites propriétaires existantes proposaient déjà des services Internet. Cependant, du fait de leur coût prohibitif, elles étaient destinées au marché des professionnels. La norme DVB-S a remédié à ce problème en standardisant une méthode d'encapsulation d'IP sur la voie aller, à destination des terminaux satellites. L'interactivité des services Internet nécessite également que les terminaux satellites soient en mesure d'émettre des données. Cette voie de retour a d'abord été supportée par une voie de retour terrestre. Une seconde étape a été la définition d'une voie de retour DVB-RCS par satellite.

Ces avancées déterminantes ne sont cependant pas suffisantes pour assurer le succès des réseaux satellites géostationnaire en tant que technologie d'accès compétitive au sein des futurs NGNs. En effet, optimisée pour la transmission de flux audio et vidéo multiplexés, la norme DVB-S n'a pas été conçue pour supporter nativement IP. De nombreuses améliorations sont nécessaires afin d'optimiser le mécanisme d'encapsulation des paquets IP et d'assurer la

continuité des protocoles de sécurité, multicast ou résolution d'adresse de niveau IP sur le segment satellite.

Par ailleurs, pour les systèmes supportant une voie de retour par satellite, leur nature fortement asymétrique reste un obstacle au déploiement de services NGN multimédias symétriques. Les technologies permettant le support de tels services sont encore trop peu avancées pour être détaillées dans la norme DVB-RCS somme toute relativement jeune. De ce fait, les solutions adoptées sont toutes propriétaires. A ce titre, il est difficile de les évaluer et l'interopérabilité entre les systèmes des différents constructeurs reste très limitée.

Afin de remédier à cet état de fait, nos travaux s'axent dans un premier temps sur la définition détaillée des mécanismes de partage des ressources partiellement spécifiés par la norme DVB-RCS. Ensuite nous incluons ces mécanismes ainsi que les dernières avancées technologiques satellites (multifaisceaux, intelligence embarquée) au sein d'une architecture de QoS satellite complète intégrant verticalement l'ensemble des mécanismes de QoS allant de la couche 2 à la couche applicative. De même, nous abordons les problématiques transverses pour notamment assurer la continuité des services IP terrestres existants ou à venir sur les réseaux satellites.

Organisation du manuscrit

Le mémoire s'articule autour de quatre chapitres.

Le premier chapitre présente un état de l'art de notre contexte d'étude. Après la présentation des nouveaux concepts propres aux réseaux NGNs, nous détaillons les principales étapes de l'évolution des architectures de QoS dans les réseaux IP au cours des dix dernières années. Les architectures de QoS NGN n'étant qu'à un stade précoce de spécification, nous nous inspirons de cette analyse afin d'établir un modèle de QoS NGN. A partir de ce modèle, nous commentons la place des technologies sans fil d'accès dans cette infrastructure NGN. L'intérêt de ce chapitre est de recenser l'ensemble des points communs et des points de rupture des réseaux NGNs avec les réseaux de communication actuels et d'en déduire les adaptations à apporter aux réseaux satellites afin de faciliter leur intégration dans le cadre des NGNs.

Le deuxième chapitre offre un panorama des dernières avancées technologiques dans les réseaux satellites. Nous étudions leur degré de maturité et nous nous attardons sur les plus prometteuses dans le cadre d'une intégration dans les réseaux NGNs. De ce fait, nous introduisons la norme DVB-S liée à la voie aller par satellite et recensons les méthodes d'encapsulation d'IP. Ensuite nous détaillons la norme DVB-RCS liée à la voie de retour par satellite en décrivant notamment les mécanismes d'encapsulation disponibles et les différentes méthodes d'accès au médium proposées au niveau 2. Enfin un point est fait sur l'intérêt des technologies multifaisceaux et l'intelligence embarquée. Après un tour d'horizon des maigres recommandations concernant la QoS dans les réseaux satellites, nous concluons sur la nécessité de concevoir une solution détaillant ces mécanismes partiellement spécifiés par les normes et les intégrer au sein d'une architecture de QoS complète.

Le chapitre 3 présente le cœur de notre contribution : une architecture de QoS pour les systèmes DVB-S/RCS. Cette architecture de QoS est basée au niveau IP sur l'implémentation d'une architecture de type DiffServ afin de répondre aux exigences de passage à l'échelle. Dans un premier temps, nous décrivons les mécanismes de QoS retenus sur la voie aller. Dans un deuxième temps, nous mettrons l'accent sur la QoS mise en œuvre sur la voie de retour, ce qui constitue la contribution la plus innovante. Nous détaillons dans un premier temps les mécanismes de QoS de niveau 2 basés sur un contrôle d'admission orienté connexion et une allocation dynamique de bande passante à la demande. Ensuite nous insistons sur l'interaction entre ces mécanismes et les classes de service implémentées au niveau IP. Dans une quatrième partie, nous décrivons la plateforme d'émulation sur laquelle cette architecture a été implémentée. Enfin, dans une cinquième partie, nous développons les dernières briques de notre architecture basées sur des solutions applicatives permettant à des applications non conscientes de

l'architecture DiffServ de bénéficier d'un traitement différencié sur le segment satellite.

Le chapitre 4 est consacré à l'évaluation des performances offertes par l'architecture de QoS spécifiée. Pour ce faire, nous caractérisons, dans un premier temps, la méthode et les outils retenus pour évaluer cette architecture. La deuxième partie expose les résultats expérimentaux obtenus et nous concluons finalement par une caractérisation de la qualité de service disponible sur le réseau satellite à travers cette architecture.

En conclusion générale de ce mémoire, nous résumons les contributions majeures de nos travaux et dégageons leurs perspectives ultérieures.

Chapitre 1 : La QoS dans les réseaux de nouvelle génération

I.1 Introduction

L'ensemble des architectures de communication est actuellement en train d'évoluer vers une infrastructure globale basée sur IP : les réseaux de nouvelle génération (NGN). Cette évolution s'explique par le fort pouvoir intégrateur d'IP qui est en mesure d'offrir un mode d'acheminement des données indépendant, d'une part, du type de technologies réseaux sous-jacentes et, d'autre part, du type de données véhiculées. L'objectif des opérateurs de télécommunications est ainsi de réaliser, à travers ces futurs NGNs, le support de multiples services (téléphonie, télévision, services Internet) au sein d'une unique infrastructure tirant parti de l'hétérogénéité des technologies réseaux.

A la frontière entre le monde des télécommunications et des réseaux, nos travaux visent à faciliter l'intégration des réseaux satellites dans ce contexte NGN en les dotant d'une architecture de QoS compatible avec celle des NGNs.

Dans un premier temps, nous allons donc réaliser une courte rétrospective sur l'évolution des architectures de télécommunications traditionnelles vers les NGNs. Dans un deuxième temps, nous allons évoquer les principales caractéristiques de ces réseaux qui n'en sont encore qu'à leurs débuts. Nous aborderons ensuite les principales évolutions des architectures de QoS IP qui vont nous permettre enfin de dresser une esquisse de l'architecture de QoS NGN et de détailler l'implication pour les réseaux d'accès de la mise en œuvre d'une QoS de bout en bout.

I.2 Les réseaux de nouvelle génération

I.2.1 Mise en garde : Terminologie

Les réseaux de nouvelle génération résultent d'un processus encore inachevé de convergence entre le monde des télécommunications et des réseaux. De ce fait, la terminologie employée pour décrire les NGNs est propre à chaque organisme de normalisation et emprunte à la fois au lexique des réseaux et des télécommunications. Ces deux disciplines partagent de plus des termes ne recouvrant pas forcément les mêmes concepts.

Afin d'éviter toute confusion, nous nous efforcerons tout au long du manuscrit de bien différencier les notions provenant du domaine des télécommunications, des satellites et des réseaux et de revisiter un certain nombre de concepts pour lesquels le vocabulaire peut être utilisé de manière contradictoire.

Ainsi nous prendrons soin de différencier les notions de « transport » ou de « plan de Transport » qui font référence dans les télécommunications à l'ensemble des technologies mises en œuvre pour le transfert des données tandis que la « couche de Transport » dans le monde des réseaux fait référence aux protocoles au dessus d'IP assurant la communication de bout en bout entre processus applicatifs (e.g. TCP, UDP).

I.2.2 L'avènement

I.2.2.1 La fin des réseaux dédiés

Jusqu'au milieu des années 80, les architectures de télécommunications suivent le paradigme d'« intégration verticale ». Les services de télécommunications traditionnels possèdent chacun leur réseau dédié optimisé pour le transport du type d'information pour lequel il a été conçu. Ainsi la voix est transportée sur les réseaux téléphoniques commutés (RTC) qui répondent alors parfaitement aux impératifs de QoS et d'interactivité. La télévision est diffusée par satellite ou par

voie hertzienne qui, de par leur nature diffusives intrinsèques, apparaît comme le mode de transmission le plus adapté. Enfin, X.25 et Internet assurent le transfert de données.

Cette vision des réseaux de communication se révèle progressivement étroite par bien des aspects, puisqu'elle entraîne des situations de « monopole naturel ». Elle se concrétise par la mise en œuvre systématique de solutions propriétaires et d'infrastructures qu'il est difficile de faire évoluer ou d'interconnecter.

Internet, bien plus qu'une simple infrastructure dédiée au transfert de donnée, va proposer une nouvelle vision des architectures de communication. L'architecture TCP/IP, découlant du modèle OSI, se révèle ainsi beaucoup plus souple et ouverte que celle proposée par l'ancienne industrie des télécommunications.

1.2.2.2 La révolution Internet

L'explosion du volume du trafic véhiculé sur Internet et l'accroissement de son hétérogénéité en terme de services, multimédias notamment, légitime définitivement l'architecture d'Internet au cours des années 90.

Basé sur le modèle TCP/IP (Application/Transport/Réseau/Liaison/Physique) et la commutation de paquet, le développement de protocoles ou de services sur Internet en est largement simplifié puisque ce modèle est ouvert, indépendant d'une architecture particulière et propose une hiérarchie protocolaire qui permet à chaque couche de s'abstraire des difficultés soulevées et résolues par la couche inférieure.

IP, en particulier, offre une connectivité en mode paquet indépendante du réseau sous-jacent permettant d'interconnecter tous types de réseau. Avec les protocoles de niveau transport (UDP/TCP), les applications disposent alors d'une interface standard pour transmettre sur un réseau IP. Graduellement des protocoles de niveau applicatif (FTP, SMTP, HTTP) vont être tour à tour standardisés permettant la prolifération des services « classiques » d'Internet (mail, Web, FTP). Finalement, l'augmentation en puissance des terminaux utilisateurs conjointement avec l'accroissement des débits et portées des réseaux a permis d'envisager non seulement la réplication, sur Internet, des services RTC ou télévisuels mais aussi le développement de nouveaux services large bande multimédias.

Cependant, malgré la prédominance actuelle des technologies d'Internet, force est de constater que la technologie IP n'est pas en mesure actuellement de supporter une telle hétérogénéité de service tout en garantissant la même qualité de service (QoS) offertes par les architectures de télécommunications traditionnelles. Internet n'offre pas de QoS. Il se contente d'un service « Best-Effort ». Une des autres limites est l'architecture chaotique des services offerts par Internet où les serveurs applicatifs sont généralement repoussés aux extrémités du réseau et où chaque nouveau service, bien que distribué, redéfinit une architecture propre non intégrée dans une architecture de service unifiée. Ainsi chaque nouveau service de l'envergure par exemple de Skype [1] ou MSN Messenger [2], aurait intérêt à s'appuyer sur des modules réutilisables de signalisation d'appel, de facturation, de gestion de profil utilisateur. Mais, pour l'instant, ces architectures de service restent fortement dédiées.

1.2.2.3 Une nouvelle infrastructure de télécommunication

Au cours des années 90, l'ensemble des acteurs du monde des télécommunications a commencé à véritablement remettre en question les architectures traditionnelles. L'ouverture à la concurrence des marchés de télécommunication, l'évolution des comportements des utilisateurs, le succès d'Internet et de la téléphonie mobile n'y sont pas étrangers. Les opérateurs vont progressivement introduire les principes fondamentaux des réseaux de nouvelle génération (NGN).

Les nouveaux défis techniques des NGNs sont alors la convergence des services (voix, vidéo et données), la convergence des réseaux autour du mode paquet, l'offre de service sans couture

d'un réseau à un autre (Handovers horizontal et vertical¹), l'ubiquité², le support de terminaux multimédias auto-adaptatifs, le déploiement de services disponibles de bout en bout accessibles depuis n'importe quel réseau d'accès et s'adaptant aux capacités des terminaux et du réseau sous-jacent.

I.2.3 L'architecture NGN

I.2.3.1 Définition et principes fondamentaux

A ce stade préliminaire, il n'existe pas encore une définition unique de la notion de NGN. Les définitions des organismes de normalisation tels que l'ETSI [3] et l'ITU-T [4] restent assez vagues et dressent une liste générale des principales caractéristiques des NGNs qui se veulent multi-réseaux, multi-services, multi-protocoles et multi-terminaux. Ces caractéristiques communes sont :

- Le « tout IP »
- Le découplage des fonctions applicatives du réseau de transport sous-jacent (Découplage Service/Transport)
- La convergence ou intégration des réseaux
- L'ubiquité
- La distribution de l'intelligence dans le réseau

A travers ces différents principes, l'objectif est de favoriser l'apparition d'acteurs tiers. Actuellement, les « opérateurs de réseaux de transport » ont tendance à se positionner sur l'ensemble de la chaîne de valeur allant des réseaux d'accès (section I.2.3.3) aux réseaux de transit en passant par les services (section I.2.3.7). Avec les NGNs, plusieurs types d'acteurs sont envisagés. La Figure 1 donne un aperçu de ces acteurs potentiels qui pourraient coexister au sein des NGNs.

L'apparition de véritables « opérateurs de contrôle » (section I.2.3.6) semble compromise du fait de la réticence des opérateurs de réseaux actuels à se séparer du contrôle des communications.

Par contre, le succès de l'ADSL et le développement des solutions « radios » (Wifi, UMTS, Satellite, Wimax) semblent indiquer que les « opérateurs de réseaux d'accès » sont véritablement voués à jouer un rôle important dans le cadre des NGNs.

Il en va de même des « fournisseurs de services » encore relativement peu développés actuellement si ce n'est dans les services de Voix sur IP. D'autres fournisseurs vont vraisemblablement profiter de l'arrivée des NGNs pour développer notamment des services de diffusion de contenus multimédias, de messagerie instantanée, d'applications (ASP Application Server Provider)...



Figure 1 : Segmentation des acteurs NGN (source : Arcome)

Nous nous basons sur ce nouveau modèle d'acteur pour introduire la terminologie

¹ La notion d'« handover » fait référence à l'ensemble des techniques mises en œuvre afin d'assurer la continuité d'un service lors du déplacement de l'utilisateur d'un réseau à un autre. Un handover horizontal s'effectuera entre réseaux reposant sur la même technologie et un handover vertical suppose le déplacement entre réseaux hétérogènes.

² La notion d'ubiquité fait référence à la nature omniprésente des NGNs.

suivante sur laquelle nous nous appuyerons tout au long du manuscrit :

- Les « opérateurs de réseaux de transit »
- Les « opérateurs de réseaux d'accès »
- Les « fournisseurs de services »

Sous le terme d'« opérateurs réseaux », nous incluons les opérateurs des réseaux d'accès et de transit.

I.2.3.2 Architecture NGN : une première décomposition

Afin de répondre aux différentes exigences citées précédemment et d'intégrer ainsi les infrastructures de télécommunication existantes avec celles à venir au sein d'une seule et unique infrastructure commune, flexible et évolutive, des impératifs ont été clairement identifiés par les organismes oeuvrant pour les NGNs (3GPP, ITU-T, ETSI, ATIS ...):

- Un cœur de réseau unique et mutualisé pour tous types de réseaux d'accès et de services
- Une architecture de cœur de réseau en 3 couches : Transport/Contrôle/Services
- Une évolution du transfert des données vers le mode paquet
- Des interfaces ouvertes et standardisées entre chaque couche afin de réaliser l'indépendance des services vis-à-vis du réseau
- Le support de technologies d'accès multiples
- Le support de terminaux multiples (modulaires, multi-mode, multimédia et adaptatifs)

Après l'évolution du cœur de réseau vers le « tout-IP », la notion la plus importante reste la décomposition en plans fonctionnels séparés par des interfaces ouvertes qui assure à la fois le passage à l'échelle et la flexibilité d'une telle architecture en offrant une facilité d'interconnexion et d'intégration de nouveaux services.

La Figure 2 offre une représentation communément admise de l'architecture des NGNs en plans horizontaux.

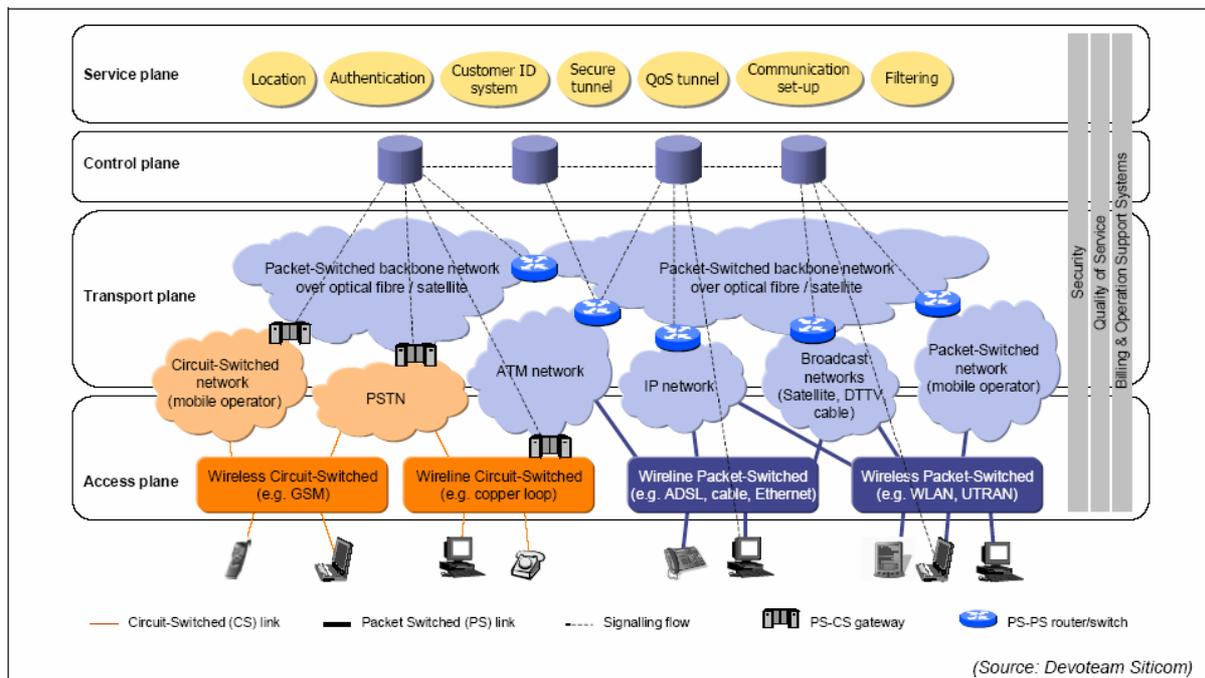


Figure 2 : Architecture générique des NGNs

Le plan de Transport regroupe l'ensemble des ressources mises en place pour assurer le transfert des données. Il gère donc l'acheminement du trafic vers sa destination avec en bordure

de réseau des « Gateways » assurant la conversion des flux de données et de la signalisation. Il dépend directement de la technologie du réseau de transport utilisé pour acheminer de proche en proche les paquets (e.g. DiffServ, IntServ, MPLS, ATM ...)

Le plan de Contrôle fait référence aux fonctions de contrôle liés à l'établissement d'une session multimédia (cf. Section I.2.3.6.2) ainsi qu'au contrôle des ressources dans le réseau de transport et l'accès au plan des Services. Elle rassemble l'ensemble des serveurs ou « softswitches » en charge des mécanismes de contrôle d'appel (Etablissement/Fermeture de session et Négociation de paramètres de session entre applications distantes) qui permettent de piloter la mise en place de ressource dans le réseau (Plan de Transport) et d'accéder au plan de Service (profil d'abonné, accès aux plates-formes de services à valeur ajoutée).

Le plan de Service regroupe les plates-formes d'exécution de service et de diffusion de contenu. Son rôle principal est de masquer la diversité technologique au client et aux fournisseurs de services. Il doit également fournir des interfaces génériques pour la création de service.

Après ces plans horizontaux, nous différencierons les réseaux d'accès des réseaux de transit. Les premiers ont pour but de concentrer le trafic des différents utilisateurs vers des équipements centraux. Les deuxièmes ont pour vocation d'acheminer des volumes de trafic importants entre quelques entités limitées. Dans la Figure 2, le plan lié à l'accès fait ainsi référence à l'ensemble des technologies permettant à l'utilisateur d'accéder aux services.

1.2.3.3 Le rôle des réseaux d'accès

Quelle que soit la définition des réseaux NGNs considérée, les réseaux d'accès sont clairement séparés du cœur de réseau. Ce dernier se doit de fournir des services multimédias aux utilisateurs indépendamment du type de réseau d'accès: WiFi, UMTS, ADSL ou satellite. Les réseaux d'accès quant à eux multiplient les technologies d'accès (réseau d'accès fixe, mobile et les réseaux locaux sans fil), évoluant constamment pour supporter de plus hauts débits (technologie ADSL, 802.11b, a, g) indispensable au support de services large bande et offrant des services adaptés à IP.

Ainsi, les réseaux mobiles 2G (GSM) et 2.5G (GPRS) ont largement contribué à habituer l'utilisateur à un usage mobile du service voix. La migration vers les technologies 3G (UMTS) a impliqué pour les opérateurs mobiles l'implémentation d'un cœur de réseau à commutation de paquet. La technologie UMTS dans sa deuxième phase représente également le premier système global entièrement normalisé avec une architecture de réseau et de services NGN. Les réseaux ADSL ou câblé ont constitué par ailleurs un des premiers exemples opérationnels de la convergence des services Voix/Vidéo/Donnée. De ce fait, les réseaux d'accès peuvent être considérés par bien des aspects comme un des éléments déclencheurs majeurs de l'évolution vers les NGNs

Enfin, par la spécificité des mécanismes régissant ces réseaux, il est important de les faire apparaître comme clairement dissociés du cœur de réseau. Cela permet de souligner que chaque réseau d'accès dispose de ses propres mécanismes d'allocation de ressource ou de signalisation fortement liés à la technologie d'accès.

Ensuite, afin de mieux illustrer les avantages d'une séparation des différentes couches du réseau de communication (Transport/Contrôle/Service), nous allons comparer la mise en œuvre de services répondant au même paradigme synchrone de type « conversationnel » [5], chacun dans leur réseau respectif : les services de communication vocale dans les Réseaux Téléphoniques Commutés (RTC) et les applications multimédias SIP dans Internet. Ces services sont les plus représentatifs dans la mesure où ils requièrent la coopération et coordination de fonctions appartenant aux trois plans de Transport, de Contrôle et de Service.

1.2.3.4 Découplage entre les couches Transport / Contrôle / Service

L'un des principaux freins à l'évolution des architectures de télécommunications traditionnelles est leur non neutralité vis-à-vis du service qu'elles offraient. Cette forte imbrication des services, du contrôle et du transport au sein d'équipements propriétaires monolithiques rend l'évolution des services très difficile puisqu'elle nécessite l'évolution de l'ensemble de l'infrastructure de contrôle de signalisation voire de transport. Ainsi les réseaux RTC sont complètement conditionnés par leur service: la communication vocale de personne à personne. Afin d'offrir un service de qualité, la technologie retenue, la commutation de circuit, suppose que les ressources soient mobilisées durant toute la durée de la communication dans l'ensemble des réseaux supportant l'appel. Ce paradigme conditionne non seulement la technologie de transport mais également la logique de conception des équipements du réseau (commutateur) et du langage de signalisation entre les équipements comme nous allons nous en rendre compte. Au contraire, Internet en offrant un mode de transfert des données non connecté s'est affranchi d'un type de service particulier et de l'architecture réseau sous-jacente. Progressivement des protocoles sont apparus afin de supporter des services de type téléphonie : RTP/RTCP [6] pour le support de médias audio et vidéo et SIP pour le contrôle d'établissement des sessions multimédias.

1.2.3.5 Plan de transport

Dans les réseaux RTC, l'architecture de transport est totalement assujettie aux communications vocales. L'infrastructure du réseau téléphonique mondial est ainsi complètement conditionnée par une granularité des ressources se basant sur des circuits de 64 kbits/s, limite choisie à l'époque qui permettait une numérisation de qualité de la voie humaine. L'ensemble des équipements de transmission du réseau mondial a été conçu et dimensionné sur ce débit de transmission. Il n'existe pas de granularité intermédiaire permettant de tirer parti des techniques de codage utilisées dans les réseaux GSM certes de moindre qualité mais n'utilisant que 8 kbits/s. Malheureusement, l'évolution des équipements dans le réseau mondial se révélerait trop coûteuse et nécessiterait un remplacement à l'échelle mondiale pour ne pas compromettre l'universalité de ce service.

Comparativement, les réseaux de données de par leur mode de transfert par paquet affichent une neutralité vis-à-vis des applications, chaque paquet IP contenant les informations suffisantes à leur acheminement entre deux terminaux. IP offre un mode de transfert homogène de bout en bout indépendant des réseaux sous-jacents et indépendant du type de données applicatives véhiculées.

1.2.3.6 Plan de contrôle

Là encore les différences entre Internet et les réseaux RTC sont majeures. Ainsi, certaines notions enchevêtrées au sein du commutateur doivent être redéfinies comme elles l'ont été par le groupe de travail TIPHON (Telecommunications and Internet Protocol Harmonization Over Networks) de l'ETSI si l'on s'en réfère à la décomposition du plan de contrôle (cf. Tableau 1) détaillée dans [7].

Tableau 1 : Décomposition TIPHON du plan de contrôle

Modèle NGN en plans	Modèle TIPHON
Plan de Service	Contrôle des services
	Contrôle d'accès aux services
Plan de Contrôle	Services de contrôle des appels
	Contrôle des « services supports » (cf. section I.2.3.6.1)
	Contrôle des médias
Plan de Transport	Services de transport
	Contrôle du réseau de transport
	Transport des flux

1.2.3.6.1 De la téléphonie classique ...

La téléphonie en répondant à un paradigme synchrone de type « conversationnel » requiert la mise en œuvre de fonctions de contrôle¹ séparées qui sont confondues dans le commutateur traditionnel avec les fonctions de transport et de service [8]: les fonctions d'accès, les fonctions d'appel, les fonctions de support et les fonctions de service.

Dans un premier temps, nous devons clarifier la notion d'appel et de « service support » (ou « Bearer Service » en anglais). L'appel se définit comme une association logique persistante entre deux ou plusieurs entités désirant communiquer et nécessitant la mise en place préalable d'un environnement de communication. Les caractéristiques des ressources associées aux transports des médias utilisés lors de cet appel sont définies par des « services supports ». Ces « services supports » dans le RTC se résument bien évidemment aux caractéristiques d'un circuit.

La mise en place de cet environnement de communication nécessite ainsi une fonction d'accès dans le réseau d'origine et de terminaison de l'appel. Dans le réseau d'origine, cette fonction permet le contrôle d'accès de l'utilisateur au réseau et le rapatriement des informations liées au profil de l'utilisateur dans son terminal (son opérateur téléphonique, son type de facturation...). Dans le réseau de terminaison de l'appel, la fonction d'accès permettait de localiser et d'identifier les capacités du terminal de son interlocuteur et de l'avertir d'un appel en cours. La fonction d'appel quant à elle supervise la mise en place de l'appel en déclenchant une ou des fonctions de service exécutant le programme d'un service de télécommunications (facturation, redirection d'appel, répondeur ...) et la fonction de support négociant et établissant les « services supports » nécessaires à la communication.

Dans le cas du RTC, et cela même malgré l'introduction des réseaux SS7 et de Réseaux Intelligents [9], le découplage entre les fonctions de contrôle d'appel et de service n'est que partiel puisque certains services sont encore rendus par les commutateurs et que le déclenchement des services intelligents s'effectue encore à partir des commutateurs. Par ailleurs, l'intelligence des terminaux utilisateurs RTC s'est longtemps réduite au strict minimum n'offrant pas la possibilité de déporter des services pourtant simples sur le terminal tel que le renvoi d'appel, ou le filtrage d'appel.

1.2.3.6.2 ...aux applications multimédias SIP

Dans le cas d'Internet, l'affranchissement des services de l'infrastructure réseau a permis nombre d'avancées impossibles dans le contexte RTC sur le plan de contrôle, notamment le support simultané de plusieurs médias de communication au cours d'un appel ou « session ». Les médias audio et vidéo ont à leur disposition de nombreux formats de compression. Cependant très peu de codecs adaptent le profil du trafic généré au type ou à l'état de charge du réseau sous-jacent et très peu d'applications permettent à l'heure actuelle de changer dynamiquement de codecs en cours de session. La session multimédia peut également incorporer des tableaux blancs, des services de messageries instantanées, de partage d'application ou d'échange de fichier comme cela est généralement le cas dans les environnements de travail coopératif.

L'environnement de communication lié à une session est de fait infiniment plus complexe que lors d'un appel téléphonique classique puisqu'il faut à la fois prendre en compte les média, les codecs, les caractéristiques des réseaux d'accès, la réservation des « services supports » nécessaires au bon fonctionnement de chaque média, le profil de l'utilisateur, les services autorisés par le fournisseur de services... L'ensemble de ces paramètres ne peut pas être pris en compte dans un contexte tel qu'Internet où la seule QoS disponible est le Best-Effort et où les applications ne sont pas conscientes des caractéristiques du réseau sous-jacent. Dans le contexte NGN par contre, l'ensemble de ces paramètres doit être pris en compte. Ainsi, dans Internet, les principaux

¹ Fonction de contrôle: fonctions exécutées afin de créer, modifier et libérer un environnement de communication nécessaire au bon déroulement d'une session multimédia (dans le cas du RTC une conversation vocale). On parle de fonction de contrôle et non d'administration dans la mesure où leur durée d'activité est liée à la durée de vie de la session.

paramètres à négocier pour mettre en œuvre une session multimédia sont :

- L'adresse IP du terminal de l'interlocuteur
- Les capacités du terminal en terme de médias et donc notamment de codecs
- Les paramètres de connexion pour chaque média (Port Source, Port Destination, Protocole de transport)

Pour ce faire, deux protocoles sont à notre disposition : SIP (Session Initiation Protocol) [10][11] et SDP (Session Description Protocol) [12]. SIP est un protocole de signalisation pour l'établissement notamment de sessions multimédias sur IP. SIP est ainsi bâti sur une architecture Client/Serveur et utilise des messages textuels transportés soit par UDP soit par TCP. Les entêtes de ces messages définissent les paramètres de routage du message et se basent sur des URIs SIP (Uniform Resource Identifier), identifiant unique de l'utilisateur. Le corps du message définit les caractéristiques des sessions à partir d'un format de description qui est généralement SDP (recommandé mais non obligatoire).

SIP réalise ainsi tout d'abord les premiers impératifs d'une signalisation de contrôle d'appel en formalisant l'établissement et la clôture d'une session entre deux terminaux et en spécifiant un protocole de négociation [13] des paramètres de session en fonction des capacités de chaque terminal (Figure 3). Pour ce faire, il supporte les fonctions élémentaires pour localiser l'interlocuteur, pour déterminer sa disponibilité (envie de participer à la communication) et ses moyens de communication (médias supportés), pour négocier un contexte de communication cohérent et gérer l'établissement de la session (ouverture et fermeture des appels).

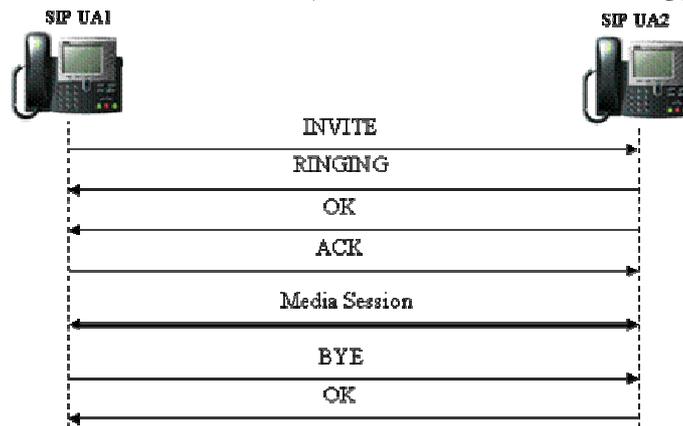


Figure 3 : Session SIP Basique

Afin de mettre en œuvre ces fonctionnalités de base, SIP ne peut se contenter que de définir des messages. Il définit en fait une véritable architecture permettant de rendre les différents services d'enregistrement, de localisation et de redirection. Les composants de cette architecture sont :

- Le User Agent (UA), User Agent Client (UAC) et User Agent Server (UAS), composants de l'application final de l'utilisateur
- Le serveur d'enregistrement (Registrar) auprès duquel l'utilisateur s'authentifie et s'enregistre en communiquant son URI SIP et son adresse IP.
- Le serveur de localisation (Location Server) qui accède à une base de donnée tenue à jour par le serveur d'enregistrement permettant de réaliser la correspondance entre SIP URI et adresse IP
- Le serveur de réacheminement (Redirection Server) aide à localiser l'adresse où l'interlocuteur pourrait être éventuellement joignable en offrant une liste d'URIs SIP alternatives.
- Le serveur mandataire (ou Proxy) relaie la signalisation SIP sans maintenir d'états liés à cette session en son sein (mode Stateless) ou bien décider de conserver des états sur le déroulement de cette session (mode Stateful) et ainsi modifier le premier message SIP (INVITE) de façon à forcer les messages SIP suivants de la session à passer par lui. Ce dernier mécanisme prend tout son sens s'il est relié à des services de facturation par exemple.

A ce stade cependant, SIP n'en était qu'à ses balbutiements puisque de nombreux services vont venir s'ajouter à la liste à partir de la 2^{ème} version des spécifications [11] qui propose des protocoles et une infrastructure stables. Les fonctionnalités avancées touchent aussi bien au niveau contrôle d'appel (propagations des authentifications, interaction avec l'ISUP (ISDN User Part) de SS7, mobilité, sécurité) qu'au niveau service (normalisation de langages de programmation de services (simples : CPL et servlets SIP, et évolués : SIP CGI), serveur de présence, messagerie instantanée)

SIP apparaît aujourd'hui comme le protocole qui va permettre l'intégration des services de voix/vidéo/donnée et la séparation formelle des couches de service, de contrôle et de transport en offrant des interfaces normalisées (protocoles et applicatives) entre ces couches. Sa généralité lui confère de plus un caractère fédérateur permettant de l'utiliser à de multiples fins tout en assurant une cohérence globale. Son ouverture, sa souplesse, et la nature distribuée de son architecture en font le protocole fédérateur des réseaux NGNs et un véritable vecteur de mise en place de modèles d'acteurs plus ouverts. Cette vision est confortée notamment par l'adoption par le 3GPP de SIP comme protocole de contrôle d'appel et d'invocation d'applications dans l'IP Multimedia SubSystem (IMS) [14] des réseaux UMTS.

A coté du contrôle d'appel et du déclenchement des services indépendants de la couche de Transport, SIP est également envisagé dans la négociation générique de « services supports » évolués dans le cadre des NGN avec la couche transport répondant aux exigences de généralité vis-à-vis des réseaux d'accès et des mécanismes de réservation de ressource associés. Ces fonctionnalités feront l'objet d'une section I.3.4.

I.2.3.7 Plan de service

La variété des services envisageables dans les NGNs est due aux multiples possibilités qu'ils offrent en terme de média, de mode de communication, de mobilité, des besoins d'adaptation au réseau d'accès et terminaux et de personnalisation. Ces services (messagerie, diffusion de contenus multimédias, services associés à la géolocalisation, l'e-commerce ...) ne pourront se déployer sans fournir aux clients des garanties de confidentialité, de sécurité et de QoS. Pour ce faire, les fournisseurs de services doivent pouvoir disposer d'outils permettant de gérer ces services (facturation, authentification, métrologie, profil des utilisateurs, états des appels). Les outils de gestion seront fournis selon deux modèles : modèle OSA/Parlay et Web Services.

Le modèle OSA/Parlay définit une architecture permettant aux plates-formes de services d'utiliser les ressources du réseau telles que le contrôle d'appel, la gestion de conférence et les informations de localisation à travers une interface standardisée masquant la complexité des réseaux. Ce modèle est préconisé pour des services de type « télécoms » nécessitant une forte coopération avec des entités de contrôle d'appel.

Le modèle complémentaire des services Web se révèle moins complexe à mettre en œuvre et correspond plutôt à des fournisseurs de services tiers. Ils reposent sur des technologies Internet (XML et SOAP) fournissant des services distribués et impliquant une forte participation des terminaux.

Nous remarquerons enfin que, pour les opérateurs de téléphonie mobile, l'IMS est en passe de devenir la première plate-forme de service unifiée qui permet une rationalisation des processus de développement et une mutualisation des infrastructures de gestion. Cette architecture adoptée et spécifiée pour les réseaux UMTS fait l'objet d'ailleurs d'un large consensus dans la mesure où le groupe TISPAN (Telecommunications and Internet converged Services and Protocols for Advanced Networking) de l'ETSI souhaite l'adapter pour la prise en compte d'accès fixes. Cette adaptation est nécessaire puisque l'IMS incorpore des fonctionnalités de contrôle propres aux réseaux UMTS.

I.2.3.8 Des services homogènes de bout en bout

Enfin la mise en place de services homogènes de bout en bout ne peut se faire sans « une vision globale et transverse [...] au niveau de la couche de contrôle et/ou de service » [15] des réseaux NGNs. Cela ne présume en aucun cas d'une homogénéisation des systèmes d'information réseau entre opérateurs réseaux. Toutefois la standardisation d'un minimum de paramètres et de protocoles homogènes entre les différents opérateurs réseaux et fournisseurs de services est nécessaire afin de permettre la mise en place de services de bout en bout. Cet aspect est plus clairement illustré dans le contexte de la QoS à la section I.4.

I.2.4 Conclusion

L'architecture NGN comme nous l'avons montré est loin d'être achevée. Héritée en partie d'Internet, cette architecture va se heurter aux mêmes problèmes, notamment celui de la QoS. Il semble cependant que le nouveau point de vue introduit par les NGNs d'une segmentation des tâches entre différents opérateurs et fournisseurs à travers différentes couches verticales (Equipements utilisateur, Réseau d'accès et Réseau de Transit) et horizontales (Transport/Contrôle/Service) permette d'envisager des solutions de QoS de bout en bout résultant de l'agrégation des solutions déjà proposées par la communauté scientifique pour l'Internet.

C'est pourquoi, avant de proposer une vision de l'architecture de QoS NGN, nous allons dans un premier temps nous référer aux principales architectures de QoS développées par l'IETF.

L'objectif est de démontrer les principales avancées qui ont permis au cours des dernières années de concevoir des architectures de QoS dynamiques:

- Un modèle hiérarchique des services
- Un provisionnement automatisé des ressources et des politiques de QoS dans le réseau
- Une architecture de renégociation dynamique des services

I.3 Architectures de QoS existantes

I.3.1 Préambule

Le fil conducteur de cet état de l'art sur les mécanismes de QoS IP existants est le suivant : l'évolution de la QoS Internet initialement centrée sur les ressources est influencée par la vision NGN de la QoS centrée, elle, sur les services.

Cette évolution est illustrée par l'importance grandissante de la signalisation de session dans le déclenchement de réservation de ressource dans les réseaux d'accès et de transit.

Quelque soit le paradigme considéré, la QoS repose toujours sur les mêmes mécanismes élémentaires (cf. section I.3.1.1) que nous allons rappeler.

I.3.1.1 Les blocs élémentaires de QoS

La mise en œuvre de la QoS au niveau IP nécessite la combinaison de mécanismes élémentaires qui sont aujourd'hui clairement identifiés et qu'il nous faut introduire brièvement afin de s'entendre avec le lecteur sur une terminologie partagée.

Pour ce faire, nous allons regrouper ces blocs élémentaires en différents plans (Données/Contrôle/Administration) en se référant à la recommandation Y.1291 définie par l'ITU-T [16] et en se référant aux différents niveaux d'abstraction définis précédemment.

Le plan de donnée comprend tous les mécanismes au niveau *élément de réseau* interagissant avec les données « utilisateur ». Ces différentes fonctionnalités sont :

- la classification
- le conditionnement qui comprend un « Shaper » et un « Policer »
- le marquage

- l'ordonnancement
- la gestion de file d'attente
- le contrôle de congestion.

Le plan de contrôle comprend l'ensemble des mécanismes qui vont permettre de configurer les ressources le long du parcours des données utilisateurs notamment lors de l'établissement de sessions multimédias. Ces mécanismes sont donc dynamiques et synchronisés avec les différentes étapes de la signalisation de session. Ces différentes fonctionnalités sont :

- le routage basé sur des informations de QoS
- le contrôle d'admission basé sur l'état des ressources disponibles
- le contrôle d'admission basé sur une politique de contrôle (« Policy Control »)
- la signalisation de QoS qui regroupe réservation de ressource, allocation de ressource, découverte et négociation de services de QoS.

Enfin le plan d'administration contient l'ensemble des mécanismes concernant les aspects de gestion et d'administration du trafic utilisateur. Il comprend ainsi la métrologie qui permet le contrôle et la supervision de la QoS dans le réseau, la gestion de contrat, la mise en correspondance des paramètres de QoS d'un service en paramètres réseau et la politique de QoS.

1.3.1.2 Trois architectures de QoS

Trois types d'architecture vont être considérées afin de rendre compte de l'évolution des solutions de QoS envisagées dans le monde IP.

La première architecture est « orientée utilisateur » et s'appuie sur l'hypothèse suivante : fournir une architecture de QoS homogène de bout en bout. Cette hypothèse implique que l'ensemble des routeurs d'Internet supporte les mêmes mécanismes de QoS qui sont configurés à l'aide d'une signalisation de QoS unique, explicite et hors-bande.

La deuxième architecture « orientée opérateurs réseaux » voit Internet comme une concaténation de domaines administrativement et technologiquement différents. Pour surmonter cette hétérogénéité de QoS, une signalisation à 2 étages va être introduite afin d'offrir de bout en bout une interface homogène de négociation de services de QoS (signalisation inter-domaine) tout en permettant à chaque domaine de gérer sa QoS de manière souveraine au sein de son réseau (Signalisation intra-domaine).

Enfin la dernière architecture est « orientée fournisseurs de services ». Elle repose également sur une architecture de QoS hétérogène au niveau Transport NGN mais s'appuie sur les signalisations de session tels que SIP pour déclencher notamment la réservation de ressources dans les réseaux.

1.3.2 Architecture « orientée utilisateur »

Cette architecture suppose que l'ensemble des routeurs d'Internet supporte des mécanismes de gestion de QoS homogènes configurables directement à partir d'un unique protocole de réservation de ressource interprétable par toutes les entités du réseau. L'architecture qui est la plus aboutie et la plus représentative de ce type de fonctionnement est l'architecture à IntServ [17][18].

1.3.2.1 IntServ

Ce modèle est principalement fondé sur une réservation de ressources par flux dans le réseau à la demande de l'application. Derrière le terme assez large de « ressource », on entend la mise en place au sein des routeurs de 4 fonctions supplémentaires : un conditionneur, un ordonnanceur, un classifieur et un contrôle d'admission. L'ensemble de ces mécanismes permet le support de 2 types de services définis par IntServ en plus du Best-Effort : le Guaranteed Service qui garantit la bande passante et un délai d'acheminement limité et le Controlled Load équivalent à un service

Best Effort dans un environnement non surchargé. Ces garanties sont obtenues de bout en bout par la concaténation de ces garanties individuelles offertes par chaque routeur traversé le long du chemin.

Afin de pouvoir configurer ces routeurs, le modèle IntServ introduit un protocole de signalisation de QoS de bout en bout, RSVP [19][20]. La réservation, déclenchée par l'application réceptrice, configure l'ensemble des routeurs sur le chemin des données en fonction du profil de trafic spécifié.

1.3.2.2 Les avantages et les limites

Cette architecture offre la gestion de la QoS de bout en bout la plus fine qui soit puisqu'elle supporte des degrés de service non prédéfinis et contrôlés qui correspondent exactement aux besoins d'un flux applicatif. Par ailleurs, elle réduit au maximum l'hétérogénéité des protocoles de signalisation à supporter en offrant un cadre de référence homogène d'un domaine à un autre. La gestion de la QoS est ainsi entièrement distribuée dans l'ensemble du réseau. Par ailleurs, son contrôle d'admission qui reposait uniquement sur la disponibilité ou non de ressource a été complété par un contrôle d'accès aux ressources basé sur des règles de contrôle [21].

Cependant l'inconvénient majeur de RSVP reste sa non « scalabilité ». En effet ce mécanisme se révèle totalement inadéquat en cœur de réseau où la concentration de flux est telle que la charge due à la signalisation induite et le nombre d'états dans le plan de contrôle et de donnée à gérer au sein des routeurs explosent [22]. Des solutions alternatives se basant sur les mêmes principes fondamentaux (architecture de QoS homogène, un unique protocole de signalisation, déclenchement de la réservation de ressource par l'utilisateur) ont déjà été proposées notamment à travers YESSIR (YEt another Sender Session Internet Reservation) [23] ou Boomerang [24] mais ses solutions, bien que mieux adaptées, supportent difficilement le passage à l'échelle.

De ce fait, un autre paradigme a été proposé par opposition à ce modèle d'architecture à service intégré pour résoudre le problème de passage à l'échelle notamment : une architecture DiffServ [25].

1.3.3 Architecture de QoS « orientée réseau »

Avant de présenter DiffServ qui est la brique de base de nos architectures de QoS « orientée réseau », nous allons préciser qu'elle constitue la première architecture à prendre en compte l'hétérogénéité d'Internet en terme d'autorité administrative et qu'elle propose des mécanismes de QoS adaptés à la vision « opérateur de réseau ». Cependant nous allons voir que cette architecture se révèle particulièrement statique et nécessite l'adjonction de nouveaux protocoles afin de s'abstraire de la complexité de la gestion des ressources du réseau. Enfin de nouveaux protocoles de signalisation devront être introduits afin de permettre une véritable interaction entre utilisateurs et opérateurs d'une part et entre les opérateurs d'autre part.

1.3.3.1 DiffServ

Afin de surmonter le problème de passage à l'échelle posé par la gestion d'une QoS par flux dans les routeurs du cœur de réseau, DiffServ offre des services par agrégats en repoussant aux extrémités du réseau le traitement par flux. Ainsi à la frontière d'un réseau, les micro-flux applicatifs sont agrégés en quelques macro-flux associés chacun à une classe de service. Les routeurs de cœur ne traitent alors plus que des agrégats limités en nombre ce qui permet de réduire considérablement le nombre d'états maintenus sur le réseau de transit et les réseaux d'opérateurs.

Cette approche s'appuie sur un certain nombre de notions que nous nous proposons de détailler dans les sections suivantes.

1.3.3.1.1 Classes de service

La détermination des classes de service passe tout d'abord par l'adoption de métriques de QoS (cf. section 1.3.3.1.1.1) qui vont nous permettre de définir des classes de trafic aux contraintes de QoS similaires (cf. section 1.3.3.1.1.2) et enfin les classes de service nécessaires (cf. section 1.3.3.1.1.3).

1.3.3.1.1.1 Les métriques de QoS

L'évaluation d'une QoS de bout en bout passe par la définition d'un ensemble de paramètres, communs aux opérateurs de réseau, permettant d'évaluer les performances de transfert des paquets IP d'un réseau et donc de l'interconnexion de ces réseaux. En se référant aux travaux du groupe de travail de l'IETF « IP Performance Metrics » ou du groupe d'étude 13 de l'ITU-T, des métriques ainsi que les méthodes de mesure ont été déjà clairement définies afin de caractériser la qualité, les performances et la fiabilité d'un service IP. La recommandation Y.1541 de l'ITU-T définit des paramètres semblables à ceux définis par l'IETF, dont les principaux sont:

- IPTD (IP Packet Transfer Delay) : Le délai de transfert d'un paquet perdu ou non entre 2 interfaces
- IPDV (IP Packet Delay Variation) : La variation de délai entre 2 paquets consécutifs
- IPLR (IP Packet Loss Ratio) : Le ratio entre le total des paquets perdus sur l'ensemble des paquets transmis
- IPER (IP Packet Error Rate) : Le ratio entre le total des paquets erronés sur l'ensemble des paquets transmis.

Afin de pouvoir évaluer la QoS perçue par l'utilisateur à travers une application de voix ou de vidéo, des paramètres complémentaires prennent un compte la caractère subjectif de perception par l'utilisateur. Ainsi la qualité d'une communication vocale peut être typiquement mesurée à travers un critère comme le MOS (Mean Opinion Score) allant de 1 (inacceptable) à 5 (excellent) obtenu à partir de tests subjectifs [26]. Le E-model de l'ITU-T [27] propose quant à lui une méthode pour lier la qualité d'une communication de voix aux performances réseaux (délai, gigue et pertes).

1.3.3.1.1.2 Classes de trafic

A partir de la définition de ces métriques, il est possible de classer les applications à partir des attentes en terme de QoS des utilisateurs. La Figure 4 issue de [28] donne une vue d'ensemble des services basée sur leur tolérance en termes de délais et de pertes. Indépendamment de l'organisme de normalisation des classes de trafic (ITU, 3GPP [29] ou TIPHON [30]), on retrouve toujours les mêmes critères de distinction :

- Type de trafic orienté connexion ou non
- Interactivité : Temps réel, interactif ou non interactif
- Critères de QoS : perte, délai et gigue

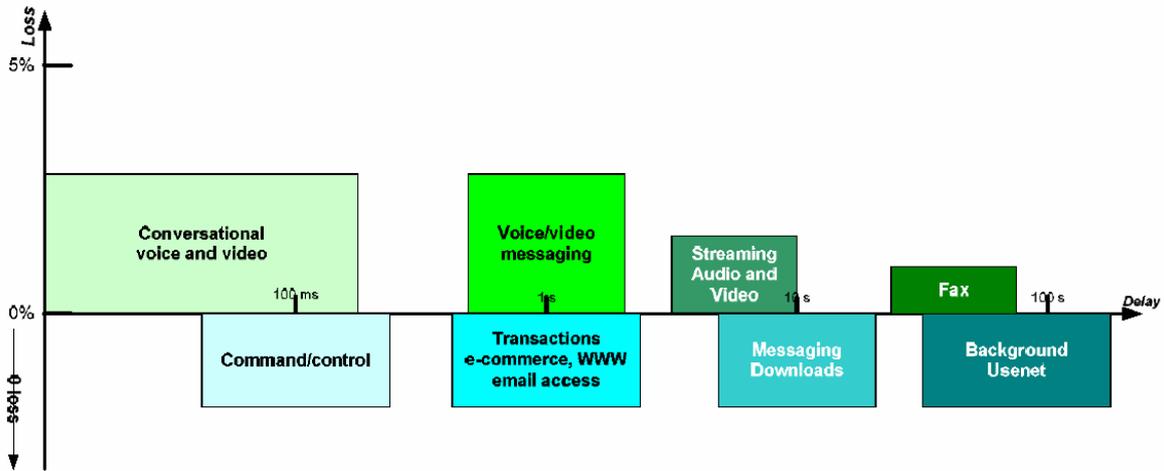


Figure 4 : ITU-T G.1010 - Besoins des applications en terme d'interactivité et de fiabilité

1.3.3.1.3 Classes de Service

De ces classes de trafic, on peut déduire autant de classes de services à implémenter dans le réseau de cœur. Il est également possible de regrouper plusieurs classes de trafic au sein d'une classe de service. Cela dépend de la politique de chaque opérateur de réseau. Cependant, des recommandations existent associant ces classes de trafic à des classes de QoS. Le Tableau 2 extrait de [31] est fourni par le groupe de travail BSM (Broadband Satellite Multimedia) de l'ETSI.

Tableau 2 : Classes de trafic BSM

BSM traffic class	Service categories	Node mechanisms	BSM resource management (see note 1)	Network techniques (informative only)	Y.1541 class	Y.1221 transfer capability	PHB (see note 2)
0	Pre-emption, emergency services, essential network services	Pre-empt any traffic that has allocated BSM bandwidth	Strict admission control with pre-emption	Strict admission control with pre-emption	N/A	N/A New Traffic Class	EF
1	Real-Time, Jitter sensitive, high interaction - Fixed size cells (VoIP)	Separate queue with preferential servicing, traffic grooming, strictly admitted	Dedicated or requested bandwidth	Constrained routing and distance	0	Dedicated Bandwidth	EF
2	Real-Time, Jitter sensitive, interactive - Variable size packets (Real Time Video)	Separate queue with preferential servicing, traffic grooming, loosely admitted	Dedicated or requested bandwidth	Less constrained routing and distances	1 (with no reference to variable size packets)	Dedicated Bandwidth	EF
3	Transaction Data, Highly Interactive, (Signalling, traffic engineering, PEPs)	Separate queue, drop priority, strictly admitted	Requested or contended bandwidth	Constrained Routing and Distance	2	N/A New Traffic Class	AF
4	Transaction Data, PEP, Interactive	Separate queue, drop priority, flow controlled	Requested or contended bandwidth	Less constrained routing and distances	3	N/A New Traffic Class	AF
5	Low Loss Only (Short Transactions, Bulk Data, Video Streaming)	Long queue, drop priority, flow controlled	Requested or contended bandwidth	Any route/path	4	N/A New Traffic Class	AF
6	Medium loss, higher delay (Traditional Applications of IP Networks)	Separate queue, flow controlled	Requested or contended bandwidth	Any route/path	5	Best Effort	Default
7	Not specified Could be used for low priority broadcast/multicast traffic or storage networks (with reliable higher layer)	Separate queue	Requested or contended bandwidth	Any route/path	N/A	Best Effort	Default

NOTE 1: The BSM resource management descriptions are informative examples only and they shall not preclude a different resource management implementation. The 3 types of resource (bandwidth) allocation are indicated as dedicated, reserved and contended.
NOTE 2: Per Hop Behaviour (PHB) are defined in RFC 2475 and RFC 2597.

La particularité de ce tableau est de recommander des mécanismes de contrôle d'admission ou de gestion de ressource adapté au type de trafic considéré ainsi que d'associer ces classes de trafic avec des classes de service. Nous noterons notamment les classes de service définies par l'ITU-T à travers la recommandation Y.1541 et les classes de services DiffServ définies à travers les PHBs (Per Hop Behavior) EF (Expedited Forwarding) [32] et AF (Assured Forwarding) [33] qui sont respectivement équivalents aux services IntServ GS et CL. Ces PHBs sont plus largement détaillés dans la section 1.3.3.1.3.

1.3.3.1.2 Domaines et SLAs

L'architecture DiffServ s'appuie sur le découpage actuel d'Internet en domaines autonomes, c'est-à-dire un regroupement de nœuds qui répondent à une seule et même autorité administrative. Au sein de ce « domaine autonome », on distingue les routeurs de bordure (routeur d'Entrée ou de Sortie qui peuvent être reliés directement à un réseau utilisateur ou à un autre domaine) et les routeurs de cœur.

Ce domaine sous l'autorité d'un opérateur de réseau offre des services. Tout client (opérateur ou particulier) qui veut en bénéficier négocie un accord avec l'opérateur concrétisé par un contrat différencié : le SLA (Service Level Agreement). Il contient des termes et des conditions techniques et non techniques. La définition du SLA par l'IETF [34] est la suivante : *“The documented result of a negotiation between a customer/consumer and a provider of a service, that specifies the levels of availability, serviceability, performance, operation or other attributes of the service”*. Ce contrat est généralement décliné en plusieurs SLS (Service Level Specification) qui constituent la partie technique du SLA : *“Specifies handling of customer's traffic by a network provider. It is negotiated between a customer and the provider, and (for example) in a DiffServ environment, defines parameters such as specific Code Points and the Per-Hop-Behavior, profile characteristics and treatment of the traffic for those Code Point”*.

Enfin la notion de contrat va de pair avec la notion d'« abonnement » particulièrement intéressante pour l'opérateur puisqu'elle lui permet de dimensionner son réseau et de vérifier la conformité des trafics injectés à partir des engagements contractés.

1.3.3.1.3 Traitement du trafic dans un domaine DiffServ

A partir de l'ensemble des SLS contractés, l'opérateur va pouvoir dimensionner son réseau et déployer les ressources adéquates au sein de son domaine par classe de service.

Au sein d'un domaine DiffServ, les routeurs de cœur traitent les paquets en fonction de la classe codée dans le champ DSCP (Differentiated Service Code Point) de l'entête IP. Ce traitement est caractérisé par un PHB qui caractérise le comportement de relayage d'un nœud DiffServ. Deux types de PHBs ont été définis par l'IETF. L'Expedited Forwarding (EF) PHB [32] correspond à la plus forte priorité et assure le transfert de flux à fortes contraintes temporelles en garantissant une bande passante avec des taux de perte, délai et gigue faible. L'Assured Forwarding PHB [33] garantit l'acheminement de certains paquets en cas de congestion. Il est scindé en quatre classes (AF1, AF2, AF3 et AF4) comprenant chacune trois niveaux de priorité. En dehors du PHB, le routeur de cœur implémente un classificateur de type « Behavior Aggregate » qui sélectionne les paquets uniquement en fonction de la valeur du champ DSCP. Ce marquage sur lequel s'appuient les routeurs de cœur est effectué par une autre catégorie de routeur DiffServ : les routeurs de bordure.

Les routeurs de périphérie, qui comprennent un métreur, un marqueur, un régulateur et un éliminateur de trafic non-conforme, sont configurés à partir des « Traffic Conditioning Block » (TCB) déduits des SLS. La tâche des routeurs périphériques est donc de conditionner le trafic entrant dans le domaine, de le classifier et le re-marquer en fonction de la politique de QoS de l'opérateur du domaine.

Ces routeurs adoptent généralement des mécanismes plus sophistiqués qu'en cœur de domaine et nécessitent notamment l'analyse de paramètres de niveau transport ou l'adresse source d'un paquet. Les routeurs de bordure sont ainsi à même de discerner au sein d'un agrégat des flux et de leur attribuer des traitements spécifiques. Des classificateurs multi-champs sont implémentés afin d'associer des classes d'applications (filtres sur des ports réservés définis par l'IANA [35]) ou d'utilisateurs (filtres sur plages d'adresse IP) à telle ou telle classe de service DiffServ ou à des mécanismes de régulation spécifiques. Des contrôles d'admission par flux peuvent être également associés afin de maintenir au sein d'une classe de service l'intégrité de QoS des flux. Ces politiques restent toutefois, à l'instar du provisionnement de ressources au sein du domaine, relativement statiques.

1.3.3.1.4 Les avantages et les limites

DiffServ propose une gestion de quelques niveaux de QoS prédéfinis par l'opérateur du domaine. Ce traitement différencié basé sur une signalisation intra-bande (champs DSCP) au sein du domaine est possible grâce à l'agrégation de trafic aux exigences similaires en bordure de ce domaine. Cette conception de l'Internet est largement plus réaliste que l'hypothèse d'une architecture de QoS homogène sous-jacente à IntServ.

Cependant, bien que cette architecture réponde aux contraintes induites par le passage à l'échelle, elle présente encore des lacunes fortement handicapantes, notamment au niveau du plan de contrôle qui est quasi inexistant.

Tout d'abord, la négociation des SLAs généralement effectuée manuellement (fax, e-mail, courrier...) reste une procédure lourde et relativement statique (mise à jour une fois par mois en général).

De plus, l'allocation des ressources n'étant généralement pas automatisée, la reconfiguration entière ou partielle des ressources au sein d'un domaine DiffServ pour supporter un nouveau services est une procédure relativement longue (des heures voire des jours). De ce fait, les ressources de ces réseaux sont généralement surdimensionnées car le réseau n'est pas en mesure de s'adapter dynamiquement à des variations de trafic trop rapides. Cependant, bien que le surdimensionnement soit une solution simple d'allocation de ressource, elle reste coûteuse et peu efficace, le facteur de surdimensionnement étant difficile à établir.

Le modèle DiffServ n'impose pas l'implémentation d'un protocole de réservation de ressource dans l'application utilisateur. Cependant les utilisateurs ne sont alors pas en mesure de modifier dynamiquement les ressources en fonction de leur besoin. De plus, l'agrégation des différents flux en agrégats peut aboutir sur une distribution inéquitable des ressources entre les flots au sein d'un même agrégat du fait de l'agressivité de quelques flux.

Enfin cette configuration manuelle implique, au niveau des routeurs périphériques, une politique statique associant des classes d'application et des utilisateurs à des classes de services ce qui n'est pas suffisant pour répondre aux changements dynamiques qui s'opèrent au sein d'un réseau. Afin d'y remédier et d'offrir une QoS de bout en bout, une solution hybride IntServ/DiffServ [36] a été spécifiée afin de combiner une gestion par flux selon le modèle IntServ dans les réseaux périphériques (réseau d'entreprise ou LAN) et une gestion par agrégat selon le modèle DiffServ dans les réseaux de transit. Cependant elle impose le support de RSVP par les applications ce qui n'est toujours pas le cas actuellement et ne résout pas le problème d'allocation dynamique des ressources dans les domaines DiffServ.

Afin de répondre à ces lacunes, de nombreuses solutions ont été proposées par l'IETF et des projets de recherche tels que Cadenus, Aquila, Tequila et Internet2-Qbone. Le point commun de l'ensemble de ces propositions est qu'elles se rejoignent sur le concept d'une architecture de gestion des ressources à deux étages [37] reposant sur une signalisation intra-domaine pour l'allocation des ressources au sein d'un même domaine et une signalisation inter-domaine pour les ressources partagées entre les domaines. Chaque opérateur est ainsi à même de gérer l'allocation des ressources dans son domaine comme il l'entend.

Dans un premier temps, nous nous intéresserons dans la section I.3.3.2 aux architectures de contrôle par politiques et au protocole COPS défini par l'IETF à travers le groupe de travail RAP (Resource Allocation Protocol) dont l'intérêt premier est la simplification de la définition et du déploiement de services au sein d'un réseau. Cette simplification de la gestion des ressources s'accompagne d'un modèle de services hiérarchique (cf. section I.3.3.3) offrant différents niveaux d'abstraction des ressources dans le réseau allant des contrats différenciés, les SLAs, jusqu'aux paramètres de configuration spécifiques correspondant à ce service dans un routeur du domaine. Dans un deuxième temps, nous étudierons les différentes extensions proposées à ce protocole pour négocier et requérir une QoS homogène de bout en bout à travers le support d'un gestionnaire centralisé de ressources inter domaines abordé en section I.3.3.5.

I.3.3.2 Contrôle de réseau à base de politique

L'architecture de contrôle par politique [38] constitue une nouvelle approche pour déployer automatiquement une politique au sein d'un domaine administratif. Une politique consiste en un ensemble de règles pour administrer, gérer et contrôler l'accès aux ressources du réseau [34]. A chaque règle correspond un jeu de conditions associées à des séries d'actions à exécuter si les conditions sont vérifiées. Afin de pouvoir réutiliser les politiques dans des domaines de gestion différents et de les déployer facilement dans le réseau, un modèle d'information général PCIM (Policy Core Information Model) est défini par le groupe de travail de l'IETF Policy. Une déclinaison hiérarchique de ce modèle est proposée afin de l'abstraire à la fois du domaine d'application (QoS, Sécurité ...), de la technologie utilisée (DiffServ, IntServ ...) et enfin de la spécificité constructeur des équipements (Cisco, Lucent ...). Ainsi, dans le cas d'une politique de QoS, un opérateur de réseau DiffServ va déduire de ses différents SLAs des règles de politiques abstraites qui seront successivement traduites en commandes génériques de configuration des routeurs puis en une séquence de commandes spécifiques au routeur considéré.

Afin de simplifier ce déploiement, cette architecture s'appuie sur différents composants détaillés dans la Figure 5.

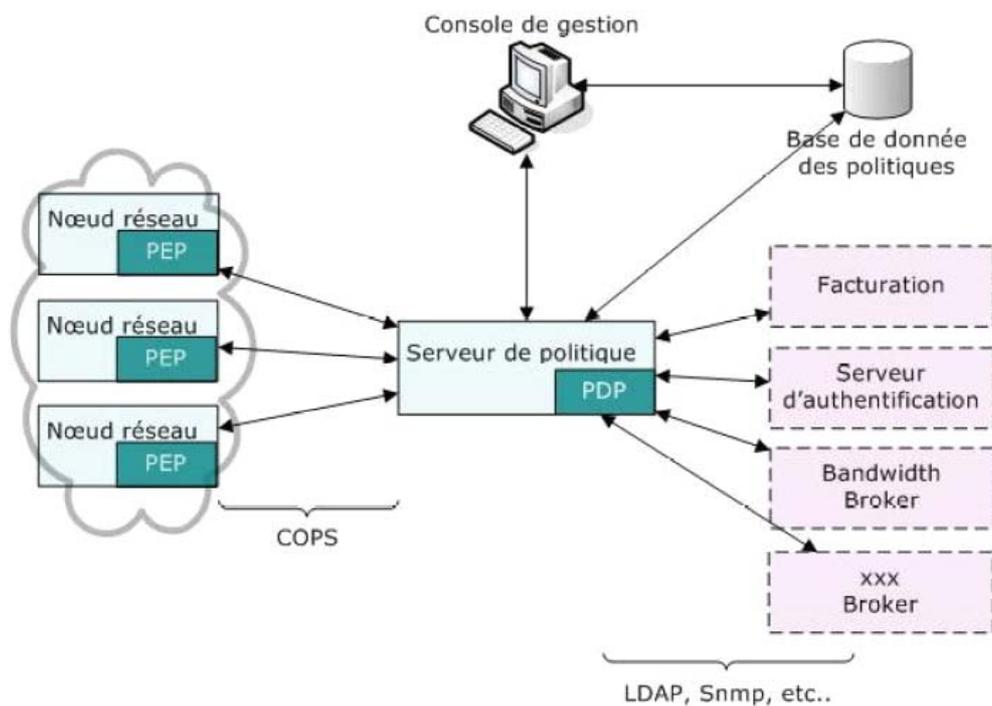


Figure 5 : Architecture de contrôle de politique

- Le PEP (Policy Enforcement Point) est l'entité logique qui applique les décisions politiques.
- Le PDP (Policy Decision Point) est l'entité logique qui décide pour elle-même et pour d'autres éléments des demandes à accepter ou non dans le réseau
- Une base de donnée des politiques (Policy Repository) stocke les règles de gestion des ressources
- Une console de gestion (Policy Management Tool) est un utilitaire permettant d'ajouter, de supprimer des nouvelles règles.
- Un gestionnaire de bande passante (Bandwidth Broker) centralise l'état des ressources en bande passante dans le réseau

Deux fonctions de contrôle distinctes sont supportées par cette architecture : l'outsourcing

[39] et le provisioning [39].

Dans le modèle « Outsourcing », des événements extérieurs (Réservation de ressource) au PEP peuvent nécessiter une décision de la part du PEP afin d'être acceptés ou rejetés. Le PEP délègue cette prise de décision au PDP en lui envoyant explicitement un message de type « Requête ». Ce dernier décide en fonction de l'état des ressources du réseau et de la politique du domaine si oui ou non cette requête peut être acceptée et répond au PEP par un message de type « Décision » où il stipule les configurations éventuelles à installer.

Dans le modèle « Provisioning », le PDP réagissant à des événements externes pourvoit les PEPs de règles que le PEP pourra appliquer localement sans solliciter préalablement le PDP. Dans ce mode, les ressources sont préallouées ce qui permet de minimiser les échanges entre le PDP et le PEP.

Enfin le protocole de communication entre le PEP et le PDP est COPS [40]. C'est un simple protocole de type requête/réponse basé sur TCP véhiculant des objets qui lui sont opaques. Chaque objet est constitué d'instances de classes dont le format et les attributs sont définis par une structure de donnée arborescente nommée la PIB (Policy Information Base). De ce fait, le protocole se révèle facilement extensible et offre à travers la PIB des classes génériques réutilisables dans différents types de politique. Ainsi la PIB DiffServ [41] définit des tables de capacités que le PEP utilise pour signifier au PDP ses capacités et la syntaxe de paramétrage des différentes fonctionnalités de traitement d'un paquet au sein d'un routeur DiffServ.

1.3.3.3 Un mécanisme d'approvisionnement dynamique des services DiffServ

Un modèle hiérarchique de services différenciés introduit dans le projet Tequila est entièrement compatible avec la philosophie de provisionnement basé sur COPS-PR [42]. Ce modèle est la clé de la séparation entre les aspects liés aux services orientés utilisateur et ceux liés aux ressources orientées réseau. La Figure 6 illustre ainsi les différentes étapes de traduction d'un SLA. Il est décliné en plusieurs SLSs comprenant des paramètres plus techniques mais toujours indépendant d'une technologie de QoS particulière (IntServ, DiffServ...). Les SLSs sont ensuite traduits en PDBs caractérisant les services DiffServ associés aux différents agrégats supportés par le domaine entre chaque couple de routeurs de bordure d'Entrée et de Sortie. Les PHBs enfin définissent les traitements appliqués dans les routeurs de cœur DiffServ. Et le dernier niveau détaille l'implémentation de ces PHBs selon des paramètres et les protocoles de configuration spécifiques au type d'équipements considérés (Cisco, Linux ...).

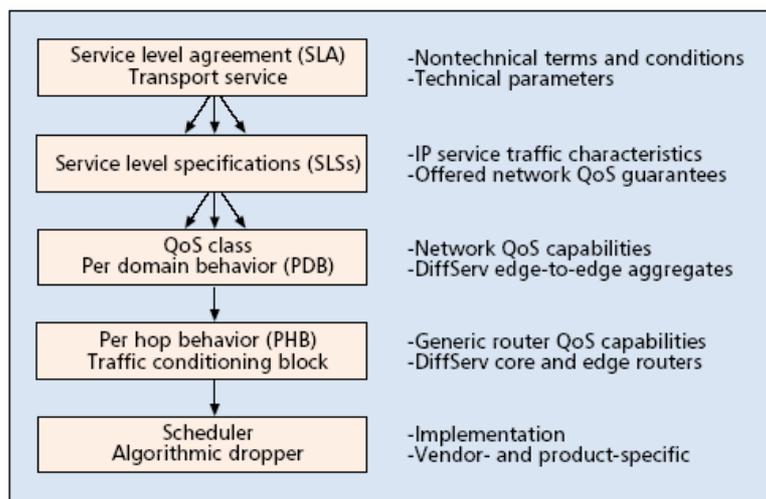


Figure 6 : Modèle hiérarchique de service différencié

Les deux couches supérieures constituent l'interface entre les clients et la couche transport IP. Les paramètres des SLS ne sont pas standardisés mais ont déjà été définis dans différents projets

de recherche comme Tequila [43], COPS-SLS [44], Eurescom [45] ou Internet2-Qbone [46]. Ces paramètres sont généralement :

- La portée (Couple Routeur d'Entrée/Sortie)
- La QoS (Service, délai, gigue, Taux de perte)
- Le profil du trafic
- Identifiant du trafic (Adresse source/destination, Port Src/Dst, Protocole de transport, DSCP)
- Identifiant du client (pour pouvoir authentifier le client)
- La période d'activité

Ensuite les notions de classes de QoS qui renvoient à la notion de PDB (Per Domain Behavior) [47] sont définies au sein de la couche liée à la QoS de niveau « réseau ». Cette couche intermédiaire assure la médiation entre la vision orientée client offerte par les SLS et les blocs élémentaires de QoS de DiffServ.

Cette décomposition permet d'affranchir les services de la technologie réseau sous-jacente, ce qui est la première étape dans la constitution d'un environnement commun de négociation lié à une description haut niveau des services disponibles dans un domaine. Nous disposons ainsi d'un mécanisme d'approvisionnement dynamique des services pour les réseaux DiffServ.

1.3.3.4 Deux Modes de gestion des services : L'« Abonnement » et l'« Invocation »

Disposant d'un mécanisme de provisionnement dynamique des ressources dans le réseau, il nous faut compléter notre architecture de QoS « orienté réseau » par un mécanisme de contrôle d'admission. Pour ce faire, le projet Tequila a pris soin cependant de distinguer deux logiques d'admission.

Ainsi l'architecture Tequila met l'accent sur la différence entre les deux processus que sont l'« abonnement à un service » et l'« invocation de service » qui font référence à deux modes de gestion des services qui opèrent à des échelles de temps différentes. L'« abonnement » permet de provisionner les profils des clients dans les bases de données liées aux services d'authentification et d'autorisation. Il gère également des services IP à long terme tels que l'accès à Internet ou la création de VPNs. Ce mode permet à l'opérateur de planifier, dimensionner et configurer son réseau à l'avance en prenant en compte les différents contrats souscrits et l'historique des trafics. L'« Invocation » est quant à elle liée à une gestion à court terme de services plus dynamiques tels que les services multimédias qui nécessitent une reconfiguration rapide du réseau. La signalisation est soit implicite, et l'utilisateur peut injecter directement son trafic dans le réseau, soit explicite et alors il fait appel un protocole de signalisation pour requérir les ressources nécessaires auprès de l'opérateur de réseau.

Se pose alors le problème de la définition d'un protocole de négociation de service dynamique permettant une allocation des ressources en adéquation avec les besoins courants du trafic.

Notre hypothèse est que les SLS sont établis statiquement et qu'ensuite l'invocation se fait dynamiquement pour les services multimédias.

1.3.3.5 Signalisation inter-domaine

Plusieurs protocoles ont ainsi été proposés afin de négocier dynamiquement des SLAs. On peut notamment citer SrNP [48], COPS-DRA [49], SIBBS [50] et COPS-SLS [44].

Ces protocoles reposent sur les mêmes principes :

- La définition de paramètres de négociation uniformes
- La définition d'un protocole de négociation
- Une entité centrale stocke l'ensemble des SLS associés aux utilisateurs ou opérateurs de domaines voisins s'étant préalablement inscrits

Ces architectures préfigurent vraisemblablement les mécanismes qui seront mis en œuvre pour offrir une QoS IP homogène sur les réseaux de cœur. Ils offrent notamment une interface de

gestion des ressources homogène, masquant l'hétérogénéité de QoS des réseaux sous-jacents, indispensable au plein épanouissement de nouveaux acteurs comme nous l'avons vu en section I.2.3.1. Cependant, malgré la maturité des architectures de QoS d'Internet, on remarquera que le « business model » des acteurs actuels d'Internet ainsi que les solutions technologiques notamment en terme de contrôle d'accès et de facturation ne sont pas encore adaptées à ces modes de négociations dynamiques de contrat de QoS entre utilisateurs et opérateurs de réseau d'accès et de cœur.

I.3.4 Une architecture de QoS « orientée session »

Très peu d'applications aujourd'hui supportent une signalisation de QoS leur permettant de signifier leurs besoins au réseau pour bénéficier d'une QoS qu'il pourrait leur offrir. Un nombre croissant d'applications intègrent aujourd'hui la signalisation de session SIP dans le but de bénéficier de l'ensemble des services de base offerts par ce nouveau protocole (localisation, négociation de média, ...) mais pas pour une signalisation orienté QoS.

I.3.4.1 Signalisation SIP, une alternative à une signalisation de QoS

SIP offre ainsi, par bien des aspects, une alternative à une signalisation de QoS. Q-SIP (cf. Figure 7) [51] constitue ainsi une des premières architectures s'appuyant sur SIP pour déclencher le provisionnement dynamique de ressources dans un domaine DiffServ. Cette allocation de ressource s'effectue sur la base des médias négociés entre terminaux SIP utilisateurs et permet ainsi à la session d'être traité prioritairement au sein du domaine DiffServ.

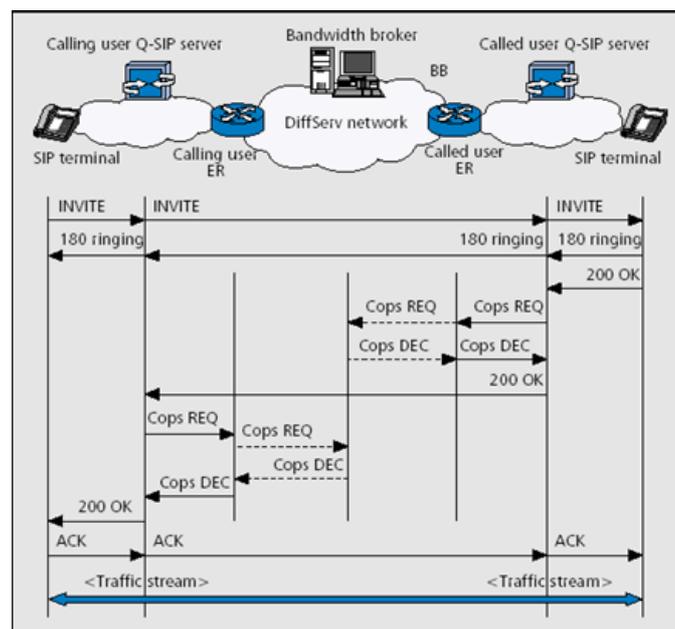


Figure 7 : Session Q-SIP basique

Cette approche souligne l'importance d'une entité relayant le trafic SIP : le proxy SIP (dans la Figure 7, il apparaît en tant que « Q-SIP server »). Il intercepte la signalisation SIP et interprète les SDP qu'elle véhicule. A travers le proxy SIP, l'opérateur de réseau est à même de connaître :

- Le début et la fin de la session multimédia (synchronisation de l'établissement de session et de l'allocation de ressource, de la fin de session et de la libération des ressources)
- Le nombre de médias impliqués et leurs caractéristiques (identification des flux et des codecs utilisés)

Les principaux avantages d'une telle architecture sont une gestion automatique de la réservation de ressources, du contrôle d'admission et de l'allocation de ressource. De plus elle est

compatible avec l'ensemble des applications SIP développées aujourd'hui car la réservation de ressources est déléguée au proxy SIP.

Cependant, ce mode d'établissement de session qualifié de « QoS Enabled » [52] déclenche la réservation de ressource parallèlement à l'établissement de la session. De ce fait, la réussite ou non de la réservation de ressource n'empêche pas l'établissement de la session. Les médias bénéficieront alors au pire d'une QoS de type « Best Effort » et, si la réservation de ressource réussit, chaque média se verra attribuer la QoS traduite à partir du SDP et requise par le proxy SIP. Ainsi, d'un côté nous allons réduire le délai nécessaire à la mise en place de la session (notamment le « Post-dial Delay »¹[53]) mais, d'un autre côté, nous allons augmenter les risques de retardement, de distorsion, ou de coupure de la voix au début de la communication (« Post-pickup Delay »²).

Afin de remédier à ces problèmes, les terminaux doivent supporter le mode d'établissement de session « QoS Assured » qui impose la réservation comme précondition à l'établissement de session.

1.3.4.2 Interaction entre signalisations de session et de QoS

Pour mettre en œuvre le mode d'établissement « QoS Assured », les terminaux doivent supporter une signalisation de QoS et une version de SIP améliorée.

Dans ce mode, les délais de réservation des ressources accroissent notamment le « post-dial delay » mais on est assuré que la communication se fera dans de bonnes conditions si notre interlocuteur accepte la communication. Cette synchronisation entre les signalisations de QoS et d'appel est possible grâce à une extension de SIP basée sur les message de type « Update » [54] qui permettent de mettre à jour les paramètres d'une session. La RFC 3312 [55] permet de notifier dans le SDP le type de réservation attendue par média et de définir ainsi si la réservation de tel ou tel média est indispensable à l'établissement de la session. De ce fait, le correspondant ne sera averti (sonnerie de téléphone) d'un appel que quand les ressources pour les médias non facultatifs à l'établissement de la session auront été réservées.

La Figure 8 illustre le scénario suivant. Un utilisateur A désire engager une visioconférence avec l'utilisateur B tout en spécifiant que la réservation de ressources pour la vidéo n'est que facultative pour l'établissement de la session. L'UA (User Agent) génère un SDP (Etape 1) et initie la session SIP en envoyant son offre SDP (Etape2). L'UA(B) génère un SDP à partir de l'intersection entre sa base de codecs et ceux spécifiés dans le SDP par l'UA(A) (Etape 3). Il le renvoie vers l'UA(A) (Etape4) et il effectue la réservation de ressource. Pour que ce scénario soit complet, considérons le cas où le point d'accès WiFi derrière lequel se trouve l'utilisateur B étant surchargé, le contrôle d'admission refuse la vidéo trop gourmande en bande passante mais alloue les ressources nécessaires à la voix (Etape 5). Les conditions à l'établissement de la session étant remplies de part et d'autre, l'UA(B) et l'UA(A) se signifient mutuellement que les réservations ont été effectuées (Etape 6). Le terminal de l'appelé UA(B) sonne (Etape 7), il envoie un message pour alerter l'appelant UA(A) (Etape 8). L'UA(B) accepte la session ce qui déclenche la fin de l'établissement de session (Etape 9) et permet la communication (Etape 10).

¹ Délai entre l'envoi d'une requête de connexion et la sonnerie du téléphone du correspondant

² Délai entre le moment où le correspondant décroche le téléphone et le moment où l'appelant entend le premier mot de la conversation.

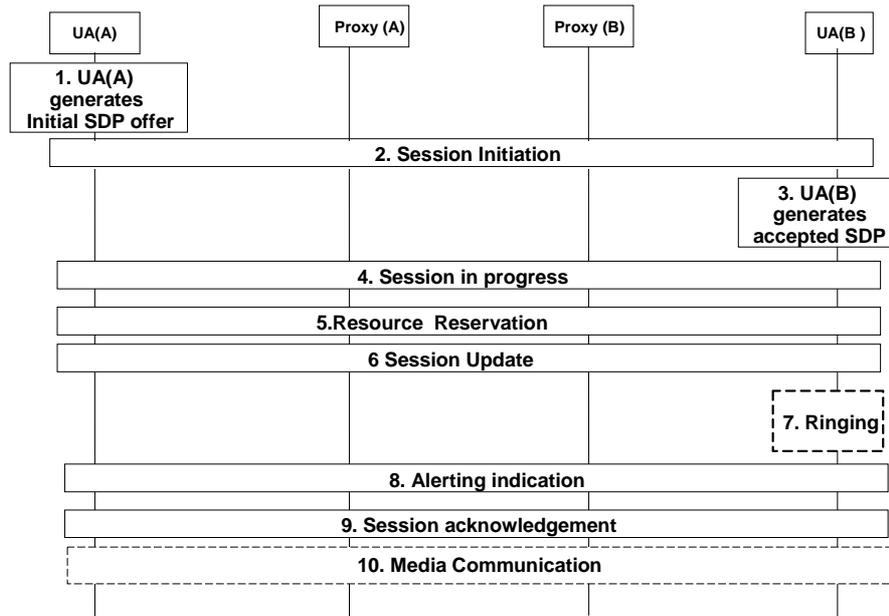


Figure 8 : Etablissement de session avec précondition

Ce scénario peut être complété par l'interaction des clients A et B avec leurs fournisseurs de services respectifs à travers les proxys A et B. Cela permet d'authentifier les clients, de vérifier si ils utilisent des services (Vidéo, Audio) auxquels ils sont inscrits, d'adapter le contenu de leur SDP ou de la signalisation SIP et de facturer le service.

Le fournisseur de services peut également vouloir s'assurer de la cohérence entre les besoins des médias autorisés au niveau session et les ressources réservées dans le plan Transport. Ce contrôle sur la réservation de ressource est effectué grâce à un jeton [56] récupéré durant la phase d'établissement par l'intermédiaire du proxy. Ce jeton est ensuite rejoué pendant la réservation de ressource permettant de réaliser la correspondance entre les quantités de ressources autorisées et celle réservées, évitant ainsi tout vol de ressource.

Enfin, afin de se prémunir d'un vol éventuel de ressource entre le moment où les ressources sont allouées et la fin de la phase d'établissement de la session (Etape 6 à 9), une « porte » peut être associée à ces ressources réservées contrôlant ainsi l'admission ou non des flux multimédias dans le réseau de cœur par exemple (si ces flux proviennent d'un réseau d'accès). Cette porte n'est activée par le proxy correspondant qu'à la réception des messages confirmant que l'appelant a bien accepté la session. L'ensemble de ces mécanismes est intégré notamment dans l'architecture UMTS standardisée par le projet 3GPP.

A travers cette section, nous avons posé les principales briques d'un modèle de QoS NGN qui est en train de se dessiner auprès des différents organismes de normalisation traditionnel :

- La segmentation du réseau NGN (réseau d'accès/cœur, concaténation de domaines administratif)
- La gestion de la QoS par niveau dans le réseau de coeur
- Le modèle hiérarchique de services
- Une gestion dynamique des ressources à deux vitesses (Abonnement et Invocation)
- Les différents modèles de requête de QoS

I.4 Architecture de QoS NGN

L'architecture de QoS des NGNs actuellement en cours de spécification dans l'ensemble des organismes de normalisation découle de cette évolution des mécanismes de QoS exposés dans la section I.3. Elle les intègre au sein d'une architecture globale non achevées mais dont les fonctionnalités et protocoles de communication principaux se dégagent progressivement. Cette section s'attachera dans un premier temps à donner une vision globale du modèle de QoS des NGNs et se concentrera ensuite sur un des véritables challenges de la QoS dans les NGNs : le support de la QoS dans les réseaux d'accès sans fil.

I.4.1 Architecture globale de QoS NGN

Si l'on se base sur les conclusions de la section I.3, afin de surmonter les hétérogénéités de politiques de QoS de chaque opérateur, de technologie de chaque domaine et d'offrir des services globaux s'appuyant sur une QoS homogène de bout en bout, le modèle de QoS NGN doit intégrer :

- Une décomposition horizontale spatiale en différents niveaux hiérarchiques de description des services regroupés en deux plans principaux Service et Transport
- Une décomposition spatiale verticale en plusieurs zones : Terminal Utilisateur, Réseau Utilisateur, Réseau d'accès et Réseau de coeur
- Une décomposition temporelle des mécanismes de provisionnement de la QoS dans le réseau: le dimensionnement du réseau (préallocation), l'invocation de service (à la demande) et les opérations de maintenance et de mesure de la QoS

Dans les NGNs, on note également que plusieurs modèles de réservation de QoS sont disponibles en fonction des capacités du terminal utilisateur et des équipements supportés par les opérateurs de service et de réseau :

- Un modèle « orienté fournisseur de service » où le fournisseur de service prend en charge la réservation de ressource auprès des opérateurs réseau
- Un modèle « orienté utilisateur » où l'utilisateur, après l'autorisation préalable du fournisseur de service (jeton d'autorisation cf. Section I.3.4.1), réserve lui-même ses ressources auprès de l'opérateur de réseau à travers un protocole de réservation de ressource de bout en bout (e.g. signalisation NSIS (Next Step In Signaling) de l'IETF [57])
- Un modèle « orienté opérateur réseau » où l'opérateur de réseau prend en charge la propagation d'opérateur en opérateur de la signalisation des besoins applicatifs d'une session.

La Figure 9 expose une vue d'ensemble simplifiée de l'ensemble des architectures de QoS proposées par l'ITU, l'ETSI, 3GPP en prenant en compte la dernière structuration globale du projet IST EuQoS [58].

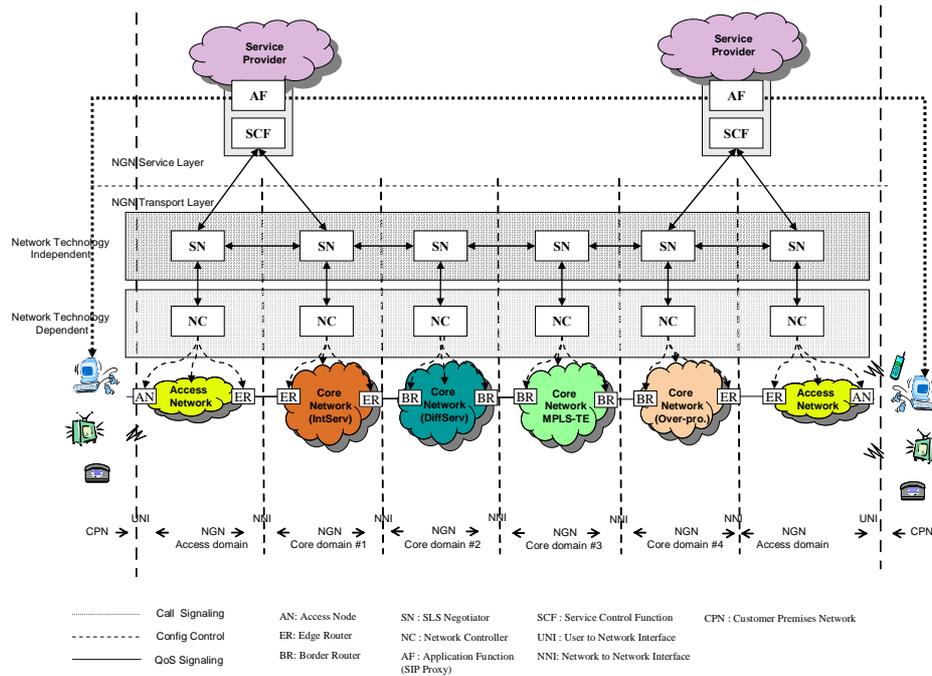


Figure 9 : Architecture globale de QoS NGN

Le point commun de ces architectures est de décomposer le plan de Transport en deux couches de gestion des ressources : l'une techno-indépendante gérée par le « SLS Negotiator » (SN) afin de pouvoir offrir au client et au fournisseur de services des interfaces génériques masquant l'hétérogénéité des réseaux et des politiques de QoS sous-jacents, l'autre techno-dépendante gérée par le « Network Controller » (NC) qui permet une gestion souveraine des ressources par l'opérateur de réseau optimisée pour le type de réseau concerné.

Au niveau Service, la signalisation SIP assure la mise en place d'un service de communication. La signalisation SIP permet ainsi d'enregistrer l'utilisateur, d'authentifier l'utilisateur et d'autoriser ou non l'utilisation de services en fonction du profil de l'utilisateur. Cette signalisation permet également de coordonner la mise en place des ressources dans les réseaux de transport impliqués. Les proxys SIP négocient avec les gestionnaires de ressources les niveaux de QoS associés à chaque média impliqué dans la session. Cette dernière fonctionnalité suppose un SDP amélioré permettant de spécifier :

- Le mode de réservation (modèle orienté « utilisateur », « fournisseur de service » ou « opérateur de réseau » et la possibilité ou d'agréger les requêtes de réservation)
- Un mode de QoS par média (QoS-Enabled ou QoS-Assured cf. I.3.4.1)

Le proxy SIP (qui réunit les fonctionnalités AF et SCF sur le schéma) est en mesure d'extraire et d'interpréter les SDPs afin de déclencher les réservations de ressources associées dans le plan Transport ainsi que le maintien de l'état de la session auprès des serveurs de facturation dans le plan Service.

La couche de Contrôle initialement présentée en section I.2.3.2 est scindée en deux parties l'une appartenant aux opérateurs de réseau et l'autre aux fournisseurs de services qui sont respectivement rattachées à la couche de Transport et la couche de Service.

Le SCF (Service Control Function) constitue une couche intermédiaire permettant aux fournisseurs de services de piloter la couche de Transport via une interface de négociation de « services supports » génériques offertes par les SNs.

Une zone d'ombre subsiste néanmoins sur le modèle de déclenchement et de propagation de la signalisation des besoins applicatifs exprimés en termes de QoS génériques dans l'ensemble des domaines traversés lors d'une communication multi-domaine.

Comme nous l'avons remarqué à la section I.2.3.2, les réseaux d'accès et les réseaux de transit ne mettent pas en œuvre les mêmes protocoles de contrôle des ressources. Alors que les réseaux de transit sont généralement surdimensionnés, leurs principales préoccupations en terme de QoS est, d'une part, l'interfaçage entre les réseaux d'accès et les fournisseurs de services et, d'autre part, les problématiques de signalisation multi-domaine (négociation dynamique de SLA entre opérateurs de cœur). Notre étude portant sur les réseaux d'accès satellite, nous laisserons de côté cette signalisation multi-domaine pour nous concentrer sur la signalisation intra-domaine et les difficultés d'offrir des services de qualité dans les réseaux d'accès basés sur des technologies sans fil.

I.4.2 QoS dans les réseaux d'accès sans fil: Problématique Multi-réseaux

Les réseaux d'accès ne peuvent se contenter des mécanismes de provisionnement de QoS de niveau IP pour offrir des services avec des garanties comparables à celle des réseaux terrestres. Ils doivent nécessairement s'appuyer sur des mécanismes de QoS de niveau 2 pour compenser leur défaut principal qui est la faible bande passante disponible.

I.4.2.1 Les caractéristiques technologiques

Malgré la hausse des débits des technologies sans fil et de leur portée au cours des dix dernières années, ces débits ne sont pas suffisants pour pouvoir se contenter de solutions de QoS de type surdimensionnement. La Figure 10 nous donne un aperçu des portées et débits des technologies sans fil actuelles. Si l'on considère également le satellite, la zone de couverture est largement plus étendue (trois satellites géostationnaires suffisent à couvrir la Terre) mais les débits sont comparables : 40Mbps en descente et 4Mbps en montée.

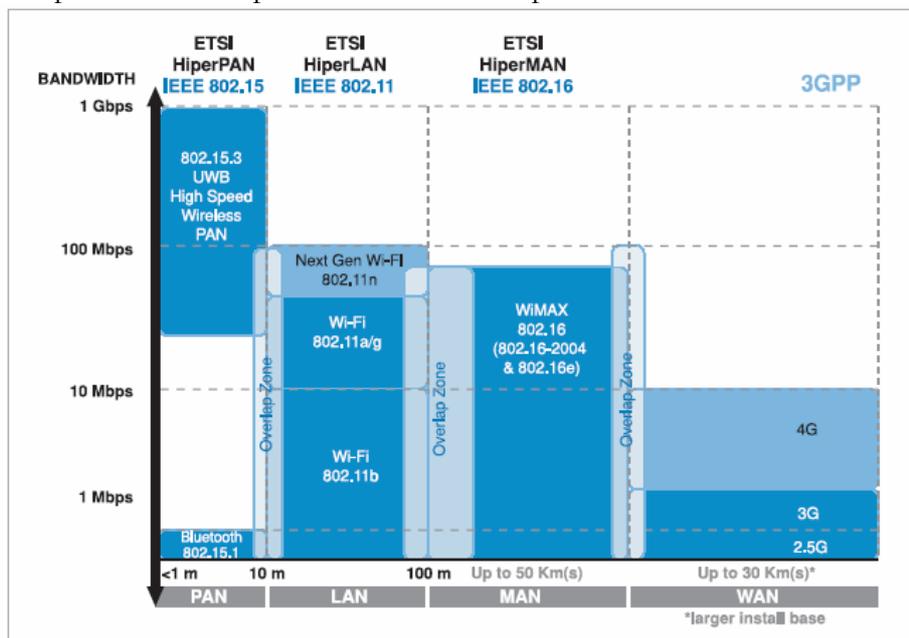


Figure 10 : Débit et portée théoriques des technologies sans fil

Chaque technologie doit être considérée à la lumière de son surcoût protocolaire, des méthodes d'accès au médium qu'elle offre, de sa zone de couverture afin de mieux cerner leur champ d'application car leur comportement n'est pas aussi fiable que celui auquel les réseaux filaires nous ont habitués.

I.4.2.1.1 L'anomalie 802.11b

Ainsi le débit maximal utile atteignable par les cartes 802.11b [59] n'excèdent pas 8Mbps du

fait d'un surcoût protocolaire non compressible. La principale raison est que le débit de transmission de 11Mbps n'est utilisé que lors de la transmission des données utiles, les entêtes des trames étant transmises pour des raisons de fiabilité à 1 Mbps (cf. Figure 11).

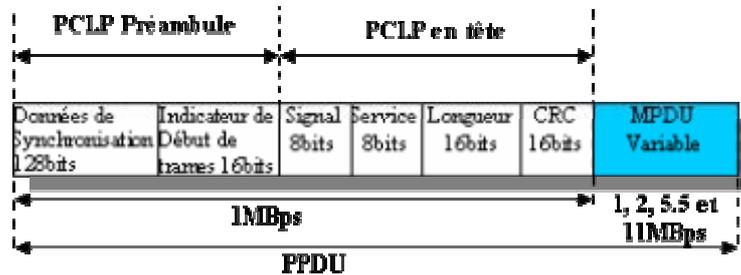


Figure 11 : Format de la trame 802.11b

Par ailleurs, la méthode d'accès principalement utilisée sur les réseaux WiFi, le DCF (Distributed Coordination Function), offre un accès équitable au médium à l'ensemble des terminaux WiFi quelque soit leur débit de transmission. Ce mode d'accès est particulièrement handicapant quand on sait que le débit de transmission des stations répond à un algorithme de variation automatique de débit calquée sur les pertes de paquets consécutives. Ainsi une station dans un environnement bruité ne bénéficie que de débits de transmission faibles (1Mbps, 2Mbps) basés sur une technique de modulation et de codage plus robuste alors que les autres à proximité du point d'accès bénéficient d'un débit « théorique » de 11Mbps. Le multi-débit couplé avec la méthode d'accès DCF a des effets particulièrement néfastes liés directement à la proportion d'hôtes lents (1Mbps) et d'hôtes rapides (11Mbps) se partageant la bande passante. La Figure 12 montre le cas extrême où une station à 11Mbps doit partager la bande passante offerte par un même point d'accès avec une station émettant à 1Mbps. Pour plus de détails, on pourra se référer aux articles suivants [60][61]

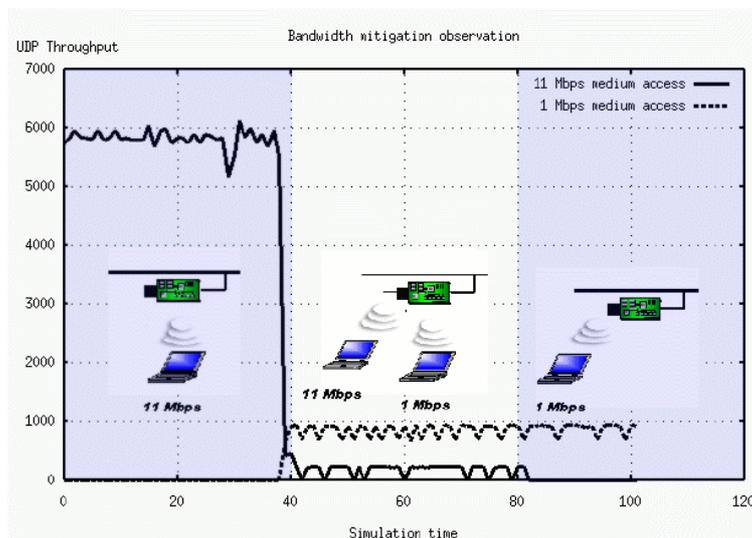


Figure 12 : Partage de bande passante WiFi entre un hôte lent et un hôte rapide

Ainsi la technologie 802.11a et g offre les mêmes travers puisque leurs débits dans la pratique ne dépassent pas 26Mbps [62] et se base sur le même mode de partage.

1.4.2.2 La QoS dans les réseaux d'accès sans fil

Le Wimax et le satellite quant à eux se basent sur une méthode de partage plus équitable du moins sur la voie de retour, c'est-à-dire de l'utilisateur à la base ou au satellite. Les ressources de

la voie de retour sont partagées en unités temporelles selon le mode TDMA (Time Division Multiple Access) allouées à la demande aux terminaux satellites ou SSs (Subscriber Stations) selon un algorithme de DAMA (Demand Assignment Multiple Access). Ces mécanismes permettent d'attribuer ainsi uniquement aux terminaux utilisateurs les ressources dont ils ont besoin et cela selon des modes d'allocation dont l'interactivité est directement liée aux besoins des applications en jeu. En adaptant la bande passante allouée à une station à sa charge et à la priorité de son trafic, on peut obtenir une gestion optimale des ressources compensant par la même la faiblesse des débits supportés. Cependant ces technologies ne sont encore qu'à leur début et aucune normalisation concernant la coopération des mécanismes de QoS de niveau IP et de niveau 2 n'a actuellement vu le jour. Il faut noter également que les algorithmes d'ordonnancement et de gestion des réservations au niveau 2 sont généralement particulièrement ouverts ce qui favorise les solutions propriétaires.

Pour les réseaux 802.11, ces mécanismes de QoS n'ont pas été inclus dans la norme au départ. La norme 802.11e [63] va bientôt standardiser le support de services différenciés grâce à des mécanismes de réservation de bande passante et à l'incorporation d'un contrôle d'admission de niveau 2. Des propositions liant 802.11e et DiffServ [64] ou IntServ [65] ont déjà été spécifiées. Cependant, il est difficile de savoir si ces mécanismes sont efficaces en environnement réel puisqu'ils n'ont été évalués qu'en simulation.

Les technologies sans fil bien qu'en plein essor n'incorporent pas toujours des mécanismes de QoS entièrement standardisés au niveau 2, ce qui limite les interactions possibles avec les couches supérieures dont notamment IP. Afin de pouvoir tirer le meilleur parti des ressources de ces réseaux, les futures architectures de QoS ne pourront se contenter de propager vers les couches basses une QoS qui ne sera que la mise en correspondance de Classes de service de niveau IP.

Des mécanismes de type Cross-Layer¹ devront être mis en place afin de pouvoir dans un premier temps remonter des informations issues des couches basses (Physique et MAC) vers les couches hautes ou inversement. Les informations des couches basses permettront aux applications de s'adapter momentanément à un environnement bruité en adoptant des codecs applicatifs moins gourmands en bande passante. Les informations des couches hautes véhiculées lors de l'établissement d'une session multimédia pourront être utilisées afin de calibrer les demandes de ressources au niveau 2.

I.5 Conclusion

Nos travaux visent à proposer une architecture de Qualité de service pour les réseaux satellites d'accès dans le cadre particulier des réseaux de nouvelle génération. Ce concept de NGN relativement jeune vise à la convergence de deux mondes, celui des télécommunications et celui des réseaux, au sein d'une même architecture globale. La difficulté d'un tel exercice est de concilier deux mondes qui ont évolué de manière disjointe avec chacun leur propre vocabulaire et leurs propres concepts.

Dans une première partie, nous avons défini les principales caractéristiques des NGNs qui sont l'adoption d'un mode d'acheminement des données de bout en bout basé sur IP et une séparation entre une couche de Service dédiée au services applicatifs et une couche de Transport regroupant l'ensemble des technologies relatives au transfert des données.

Fort de ces nouveaux concepts, nous avons abordé ensuite un des principaux verrous technologiques des NGNs : la Qualité de Service. Aucun modèle standard n'existant à l'heure actuelle, nous avons détaillé l'évolution des architectures de QoS dans le monde IP en les remplaçant dans un contexte NGN. Un modèle global de QoS NGN a été finalement esquissé.

¹ La notion de Cross-layer fait référence à l'ensemble des méthodes visant à l'optimisation des performances d'un protocole basée sur l'échange d'informations inter-couches. Ainsi une pile de protocole en bénéficiant d'informations de la couche physique par exemple est à même de s'adapter plus rapidement et plus efficacement aux conditions de transmission.

Dans ce modèle, nous avons opéré une nette distinction entre les réseaux d'accès et les réseaux de cœur en terme de gestion de QoS

La dernière partie de ce chapitre souligne l'impossibilité de se contenter d'une gestion de la QoS à un niveau aussi grossier qu'IP dans les réseaux d'accès sans fil où les ressources sont plus limitées que dans les réseaux ADSL ou cablés. L'accent est alors mis sur les mécanismes de gestion de QoS de niveau 2 disponibles dans ces types de réseaux qui permettent d'optimiser, s'ils sont intelligemment couplés avec IP, l'utilisation des ressources.

Les réseaux satellites géostationnaires font partie de ces technologies d'accès qui pourraient apporter une forte plus value aux réseaux NGNs en offrant une connectivité large bande et des services de QoS différenciées dans des zones géographiques dépourvues d'infrastructures terrestres. Ces réseaux sont l'objet de notre deuxième chapitre qui vise à familiariser les lecteurs avec le système de communication qu'est le satellite. Longtemps basés sur des solutions propriétaires, ces systèmes adoptent enfin des standards ouverts encore largement méconnus de la communauté réseau et que nous nous proposons de détailler dans le chapitre 2.

Chapitre 2 : QoS dans les Réseaux Satellite de Nouvelle Génération

II.1 Les satellites, la réponse à la fracture numérique ?

Actuellement en Europe la technologie prédominante pour fournir des services large bande est la technologie DSL (ADSL2+ et VDSL). Peu coûteuse car s'appuyant notamment sur une boucle locale préexistante, son déploiement ne se révèle rentable toutefois que dans les zones à forte densité¹, privilégiant ainsi les pays et régions à forte population urbaine (Grande-Bretagne, Suède, France ...) aux dépens des pays ou régions à forte population rurale (République Tchèque, Pologne, Grèce ...). Ainsi à l'horizon 2013, les services terrestres large bande resteraient inaccessibles pour près de 13 à 20 millions de consommateurs potentiels dans l'Europe. Le rapport [66] estime que le satellite serait la solution optimale pour près de 7 millions d'utilisateurs tandis que le reste serait servi par des infrastructures basées sur le Wimax. Ainsi la pénétration des services large bande si elle ne se base que sur la technologie DSL sera fortement dépendante des facteurs géographiques (densité de population, topographie, richesse). Dans ces zones à faible infrastructure terrestre, il apparaît évident que la fracture numérique géographique ne peut être résorbée, dans le court terme, que grâce aux technologies sans fil. Cependant peu d'entre elles sont en mesure à l'heure actuelle d'offrir un accès large bande avec un rayon de couverture suffisamment grand sans un déploiement préalable d'infrastructures coûteuses. La seule technologie à même de pouvoir réaliser cet exploit à court terme est la technologie satellite éventuellement couplée à des réseaux sans fil locaux.

II.1.1 Les avantages intrinsèques

Les satellites géostationnaires (GEO) quelle que soit leur utilisation présentent toujours les avantages intrinsèques suivants :

- Une infrastructure terrestre minimale
- Une zone de couverture immédiate d'un ou plusieurs pays
- Un déploiement simple et rapide des terminaux.

Alors que les réseaux terrestres sont déployés progressivement et doivent maintenir une continuité physique, les réseaux satellites ne nécessitent pas la création d'infrastructure dans le domaine public. Contrairement aux réseaux terrestres, ils ne sont pas contraints par les frontières nationales. La dimension mondiale des principaux opérateurs satellites suffit à nous en convaincre. De plus, il est possible d'installer rapidement des terminaux n'importe où dans la zone de couverture ce qui est un avantage incontestable notamment dans les situations de crise comme les catastrophes naturelles. Enfin, ces infrastructures sont moins sujettes aux catastrophes naturelles et il est plus facile de préserver l'intégrité d'un réseau satellite qu'un réseau terrestre, d'où des coûts de maintenance réduits.

Nombres de services à l'heure actuelle exploitent la capacité de diffusion naturelle des satellites. Ils offrent notamment un moyen d'interconnexion avec les services terrestres à travers les stations embarquées dans les voitures, les bateaux ou les avions. Ils permettent d'offrir une connectivité (accès Internet notamment, ou téléphonie) dans les zones géographiques reculées sans infrastructure terrestre (océan, zone de catastrophes naturelles). Ils ont permis de mettre en œuvre des services de téléphonie longue distance et se révèlent aujourd'hui particulièrement adaptés aux services de diffusion ou de communications multi-points tels que la télévision, la

¹ Au-delà de 5km des répartiteurs, la connexion de communautés isolées par l'ADSL n'est considérée comme commercialement viable que pour un nombre minimum de 100 abonnés environ ce qui correspond, avec un taux de pénétration approximatif de 25% et une moyenne de 2.45 personnes par foyer, à des villages d'au moins 1000 habitants.

radio, la vidéo à la demande (VoD) et le *push* de données.

II.1.2 Les avancées récentes

Bien que les réseaux satellites géostationnaires aient longtemps été dédiés aux services de diffusions, de nombreux efforts ont été réalisés par la communauté satellite pour en faire une technologie complémentaire aux infrastructures terrestres pour le support de services interactifs large bande dans les zones non couvertes. Ces avancées visent toutes à pallier les inconvénients majeurs des réseaux satellites qui sont :

- Les coûts élevés des équipements et des satellites
- Les temps de latence (un délai de propagation aller-retour incompressible de 500ms dans le cadre d'une communication directe)
- Leur faible et coûteuse bande passante.

L'une des premières avancées concrètes est l'adoption et la spécification par la communauté satellite d'interfaces et de standards ouverts afin de réduire significativement les coûts des équipements et d'offrir à des prix raisonnables des services aux utilisateurs en favorisant la concurrence entre opérateurs, fournisseurs de contenu et fabricants. Cette tendance est confirmée notamment par le succès de la norme DVB-S qui est actuellement le format préféré de diffusion digitale à travers l'Europe et la banalisation des terminaux récepteurs. La standardisation d'une voie de retour par satellite, DVB-RCS (Digital Video Broadcast Return Channel Via Satellite) [67], ainsi que les efforts d'interopérabilité des équipements réalisés par le groupe SatLabs s'inscrivent directement dans cette tendance.

En dotant les terminaux satellite d'une voie de retour par satellite, DVB-RCS affranchit complètement les systèmes satellites d'une voix de retour terrestre les rendant à même de supporter ainsi un accès interactif à des services multimédias large bande. De nombreuses et récentes avancées technologiques pourraient encore accroître le potentiel des systèmes satellites bidirectionnels en tant que réseaux d'accès. Ainsi l'exploitation de la bande Ka (27 à 40 GHz) jusqu'alors fortement tributaire des conditions météorologiques permettrait aux réseaux satellites non seulement de bénéficier de canaux plus larges (en comparaison avec la Bande Ku complètement engorgée) mais également de diminuer la taille des terminaux utilisateurs et des antennes. Ce gain est possible par la réduction de la couverture des faisceaux. Le recours à une couverture multifaisceaux dans un cadre transparent ou régénératif permet un gain conséquent en capacité du fait de la réutilisation de fréquence alors possible. Couplée à une charge utile régénérative, la couverture multifaisceaux permet une plus grande souplesse d'interconnexion pour une utilisation optimale des ressources. Enfin, la norme DVB-S2 [68], successeur de DVB-S, propose des techniques de codage et de transmission plus adaptées aux services point-à-point exigés par les nouvelles applications multimédias avec une amélioration de l'utilisation des capacités d'un transpondeur variant entre 100 et 200%. Ces nombreuses évolutions laissent présager une place prépondérante des réseaux satellites dans les futures technologies d'accès.

Avant de détailler les normes DVB-S et DVB-RCS sur lesquelles les architectures satellites bidirectionnelles classiques s'appuient, nous allons détailler les principaux composants et fonctionnalités de ces architectures

II.1.3 Les systèmes satellites d'accès bidirectionnels

II.1.3.1 *Vue d'ensemble*

Les systèmes satellites multimédias à large bande, quelque soit l'implémentation considérée, se décomposent en trois segments distincts :

- Le segment utilisateur qui comprend les *terminaux satellites* dotés d'une voie de retour (Return Channel Satellite Terminal : ST) intégrant des fonctionnalités d'accès au satellite et assurant

- l'interconnexion entre les équipements utilisateurs et le réseau satellite
- Le segment spatial qui est composé d'un ou plusieurs *satellites*
 - Le segment opérateur constitué d'un ou plusieurs centres de contrôle ou *Network Control Centers* (NCC) et de passerelles ou *Gateways* (GW) s'interconnectant avec les réseaux de cœur terrestres.

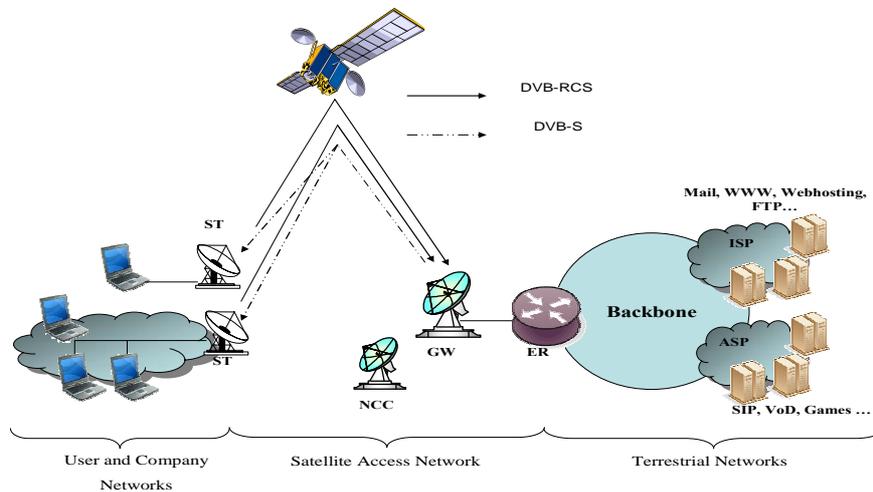


Figure 13 : infrastructure d'accès satellite bidirectionnel basique

La Figure 13 nous présente un réseau d'accès satellite classique étoilé transparent et son intégration au sein d'une infrastructure globale comme Internet.

La GW centralise l'ensemble du trafic dans le réseau satellite. Toutes les connexions issues des STs sont à destination de la GW, d'où une topologie étoilée. La GW émet en DVB-S sur la voie aller (GW → ST) et reçoit en DVB-RCS sur la voie retour (ST → GW). La GW concentre ainsi sur la voie aller le trafic provenant des ISPs (Internet Service Providers) et est doté de deux modules d'administration pour gérer les ressources satellites, le NMC (Network Management Center) et le NCC (Network Control Center). Ces entités sont responsables de l'administration et du contrôle de l'accès aux ressources satellites par les STs. Leur rôle respectif est plus clairement détaillé à la section III.3.2.2. Ici, ces entités sont incluses dans la GW. Dans une configuration régénérative, elles sont découplées de la GW et réunies au sein d'une unique entité.

Les STs se comportent comme des routeurs d'accès au segment satellite et concentrent le trafic utilisateur sur la voie retour. Ils émettent en DVB-RCS et reçoivent en DVB-S. Ils interagissent avec les modules d'administration de la GW afin d'obtenir leurs paramètres de configuration et leurs ressources satellites.

II.1.3.2 *Connectivités d'accès et maillée*

Deux types d'interactions sont à envisager dans ces réseaux :

- La *connectivité d'accès* (GW-ST) qui permet aux utilisateurs connectés à un ST d'accéder à Internet (ou d'autres réseaux) via la GW
- La *connectivité maillée* (ST- ST) qui permet aux utilisateurs connectés à des STs appartenant à un même réseau satellite de communiquer soit par l'intermédiaire de la GW soit directement de ST à ST à travers le satellite

De nombreuses configurations sont possibles pour supporter ces connectivités. Le groupe de travail BSM (Broadband Satellite Multimedia) de l'ETSI distingue ainsi plusieurs familles de systèmes satellites [69] combinant une topologie *maillée* ou *étoilée* définis dans [70] et un satellite avec une charge utile *transparente* ou *régénérative*.

II.1.3.3 Topologies étoilée et maillée

L'architecture maillée supporte ainsi des communications directes entre utilisateurs connectés à des STs appartenant au même réseau satellite ce qui se concrétise généralement par des échanges symétriques. Dans l'architecture étoilée, les échanges se font tous par l'intermédiaire d'une station centrale, la GW, et sont donc par nature asymétriques avec des débits plus importants de la GW vers les STs. Toutefois cette architecture étoilée peut également supporter la connectivité maillée en établissant deux liaisons indirectes entre les STs via la GW.

II.1.3.4 Satellite Transparent ou régénératif

L'accès à Internet s'effectue à travers 2 voies :

- La *voie aller* provenant d'un fournisseur de services ou d'accès Internet (et donc de la GW) à destination de l'utilisateur qui tire parti des capacités de diffusion du satellite vers plusieurs STs.
- La *voie retour*, des STs à la GW, sur laquelle le satellite présente alors une capacité de collecte captant les informations provenant de différents STs.

Les contraintes de partage des ressources ne sont pas les mêmes, les ressources de la voie aller étant monopolisées par la GW tandis que les ressources sur la voie de retour sont partagées entre autant de STs. Ainsi l'ETSI a spécifié deux normes l'une pour la voie aller qui est le DVB-S et une pour la voie retour le DVB-RCS. L'intérêt d'un satellite transparent ou régénératif prend alors tout son sens puisque, selon la charge utile embarquée dans le satellite, la conversion des signaux DVB-RCS en signaux DVB-S a lieu soit au sol dans la GW pour une charge utile transparente, soit dans le satellite pour une charge utile régénérative. Les deux cas de figure sont illustrés dans la Figure 14.

Ainsi, avec un satellite transparent, la communication inter-STs nécessite un « double bond » satellite alors qu'avec un satellite régénératif, elle ne nécessite qu'un « simple bond ».

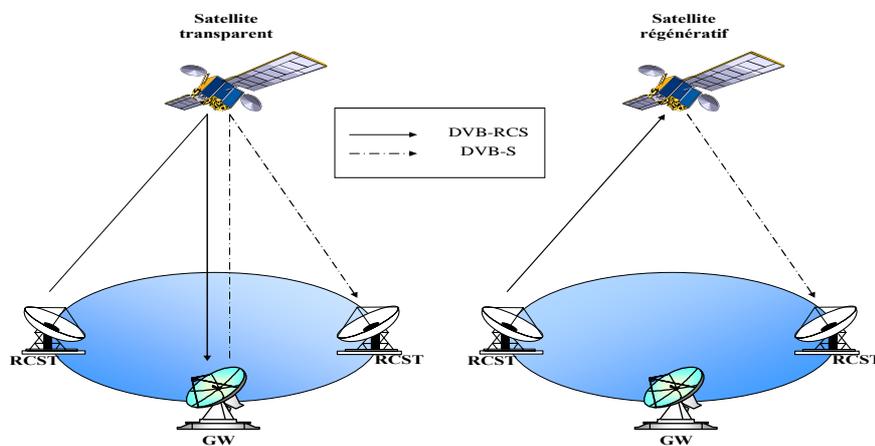


Figure 14 : Communication inter-ST avec satellite transparent et régénératif

La norme DVB-RCS précise que les *liaisons montantes* (à destination du satellite) et les *liaisons descendantes* (en provenance du satellite) doivent utiliser des fréquences différentes.

Ainsi, en mode transparent, le transpondeur¹ d'un satellite se comporte comme un simple répéteur qui reçoit sur un lien montant un canal déterminé, convertit sa fréquence et réémet ce

¹ Transpondeur : Répéteur embarqué dans le satellite qui reçoit sur un canal déterminé les signaux d'une liaison ascendante, les convertit à la fréquence de la liaison descendante et le rediffuse vers la liaison descendante. Les transpondeurs sont associés à une ou plusieurs antennes d'émission qui, par leur forme et leur orientation, définissent la zone de couverture du faisceau émis.

canal vers un lien descendant. L'amplitude et la fréquence de la porteuse sont modifiées mais la modulation et la forme du spectre restent inchangées.

En mode régénératif, le transpondeur comportent les mêmes fonctionnalités plus un traitement à bord qui inclut un traitement numérique du signal permettant le démultiplexage/remultiplexage des flux au sein du satellite. Cette capacité de traitement à bord du satellite est également qualifiée d'intelligence embarquée ou OBP (On Board Processing). On parle également de charge régénérative ou de satellite régénératif. En plus d'une réduction du délai de propagation entre deux STs, cette technologie couplée avec l'utilisation de multifaisceaux permet d'optimiser la gestion des ressources du satellite.

L'interactivité des communications inter-STs est certes accrue mais elle se fait au détriment de la simplicité et du coût du satellite.

II.1.3.5 *Satellite monofaisceau et multispot*

Chaque canal peut être défini par sa zone d'émission et sa zone de réception. Ces différentes zones permettent d'offrir un traitement similaire à un groupe de terminaux satellites qui partagent un même service. Un spot délimite ainsi une zone de couverture généralement associée à un ou plusieurs transpondeurs. La plupart des satellites traditionnels sont généralement monofaisceaux ce qui se traduit par des zones de couverture particulièrement larges. L'utilisation de plusieurs faisceaux de faible ouverture permet de séparer la couverture globale du satellite en plusieurs zones plus petites. La taille des spots étant réduite, le signal est plus concentré et son gain est plus élevé. Ainsi pour une taille d'antenne équivalente, le terminal bénéficiera d'un débit plus conséquent. Par ailleurs, le support de plusieurs spots permet une réutilisation des fréquences d'un spot à un autre et un gain non négligeable en ce qui concerne les messages de point à point qui au lieu d'être transmis vers un spot particulièrement large seront orientés directement sur le spot correspondant sauvegardant les ressources des autres spots. L'interconnexion entre spots dans un cadre régénératif est d'autant plus intéressante et efficace qu'elle permet de combiner différents flux montants provenant de différents spots montants dans les différents spots descendants cibles.

Des satellites implémentent déjà des faisceaux restreints à l'instar d'IPStar [71] ou Astra 1H composé de 32 transpondeurs et de 8 faisceaux en bande Ka. Les systèmes SkyplexNet/Hot Bird 6 [72] ainsi qu'Amerhis supportent des capacités régénératives couplées à des techniques multifaisceaux. La Figure 15 donne un aperçu d'un système satellite bidirectionnel intégrant les techniques multifaisceaux et une charge utile régénérative.

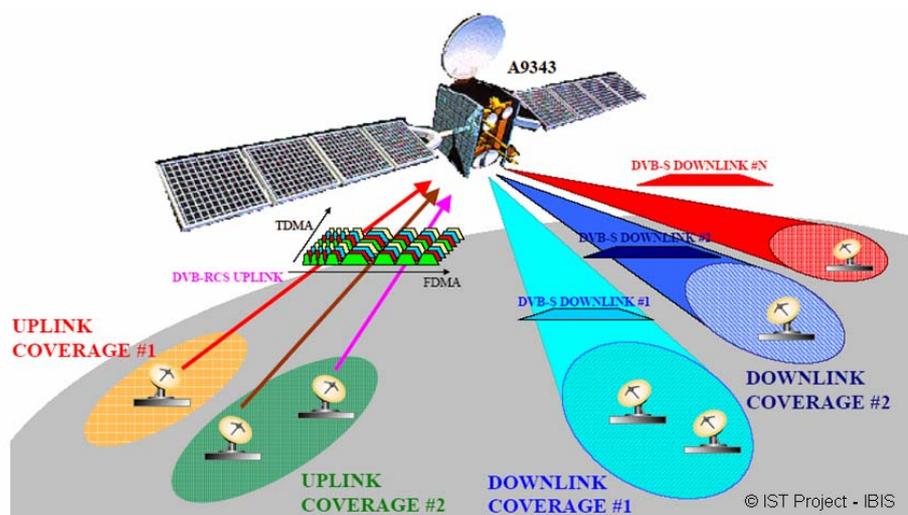


Figure 15 : Satellite bidirectionnel multi-spots régénératif

Après avoir présenté les composants et concepts fondamentaux des systèmes satellites bidirectionnels, nous allons pouvoir nous intéresser dans le détail aux normes DVB-S et DVB-RCS qui définissent des mécanismes de niveau physique, liaison et réseau. Afin de mieux appréhender les difficultés de l'intégration des réseaux satellites d'accès dans les réseaux NGNs, nous nous concentrerons sur les interactions avec IP envisagées dans chaque norme.

II.2 La norme DVB-S et support d'IP

Initié en 1993, le projet européen international DVB a publié, à la fin des années 90, une famille de spécifications de transmission digitale basée sur un standard permettant de transmettre de la vidéo et du son sous forme numérique dans un format compressé¹ développé par le groupe MPEG. Ce format est le MPEG-2 [73] qui définit à la fois une norme de compression et des méthodes de multiplexage. DVB fut décliné en autant de spécifications que de médias de transmission, les trois principaux standards étant spécifiés pour le câble (DVB-C), le satellite (DVB-S) et le terrestre (DVB-T).

Le DVB-S définit ainsi le mode de transmission sur un lien unidirectionnel satellite de données audio et vidéo vers les terminaux équipés de carte DVB-S. Des données autres que des flux multimédias sont également supportées. Les systèmes DVB transportant des paquets IP sont qualifiés de systèmes DVB de diffusion de donnée (DVB Data Broadcasting) [74][75]. Comparativement à l'architecture bidirectionnelle présentée dans la section II.1.3, l'architecture classique d'un système DVB-S dédié aux services télévisuels s'appuie sur un satellite transparent monofaisceau et une GW générant un unique multiplexe occupant la largeur de bande complète du transpondeur, les terminaux satellites se contentant de recevoir.

II.2.1 La norme DVB-S

Le succès de cette norme repose sur :

- la simplicité de la GW qui centralise les données (ensemble des programmes à diffuser) et la signalisation. Ces dernières sont multiplexés par la GW sur une unique porteuse occupant l'ensemble de la largeur de bande offerte par le transpondeur. La GW émet à débit constant utilisant des bits de bourrage si besoin en est
- La rapidité et la facilité d'installation des terminaux satellites
- Les capacités d'évolution offertes par la généricité des conteneurs MPEG2-TS (à inclure flux vidéo, audio et tous types de données privées).

II.2.1.1 Les flux de transport MPEG2-TS

Quel que soit le type de données émis par la Gateway à destination des terminaux satellites (son, vidéo ou donnée), elles sont encapsulées dans des paquets MPEG2-TS de taille fixe (188 octets) dont la séquence crée un flux de transport MPEG2-TS (Transport Stream). On parle également de multiplexe puisqu'il regroupe plusieurs canaux de données au sein d'une seule voie de communication. Plusieurs méthodes, détaillées en II.2.1.3, sont disponibles afin d'encapsuler différents types de données. Nous allons nous intéresser ici à la diffusion de flux vidéo et audio qui a fait le succès du DVB-S dans le cadre de la télévision numérique. La Figure 16 illustre les éléments participant au multiplexage des différents programmes.

Imaginons un bouquet satellite offrant différents programmes télévisuels à l'abonné. Chaque programme est constitué d'une source vidéo, d'une ou plusieurs sources audio et de données (pour les sous-titres par exemple) synchronisées sur la même horloge de référence (Program Clock Reference). Chaque source est compressée séparément et donne un signal brut en sortie de compression que l'on nomme flux élémentaire (ES : Elementary Stream). Ces flux élémentaires audio et vidéo continus sont ensuite encapsulés en paquets (PES : Packet Elementary Stream)

¹ taux de compression allant de 52 :1 pour le MPEG1 en 1991 à 200 :1 pour le MPEG-2

selon une base temporelle commune. Chaque paquet contient ainsi une estampille temporelle et un identificateur de flux (audio ou vidéo). La taille de ce paquet dépend du type d'application et peut atteindre plusieurs centaines de kilo-octets. L'ensemble des programmes du bouquet satellite décomposé en autant de PESs est enfin multiplexé au sein d'un flux de transport MPEG2-TS. Les paquets constituant ce multiplexe sont des conteneurs MPEG2 indépendants du type de trafic véhiculé (source native MPEG2, ou IP). La méthode d'encapsulation implémentée est le Data Streaming. Selon ce principe, les PES sont fragmentés et insérés dans les paquets MPEG2-TS. Un paquet PES devant toujours coïncider avec le premier octet de la charge utile d'un paquet MPEG2-TS, des bits de bourrage sont alors nécessaires.

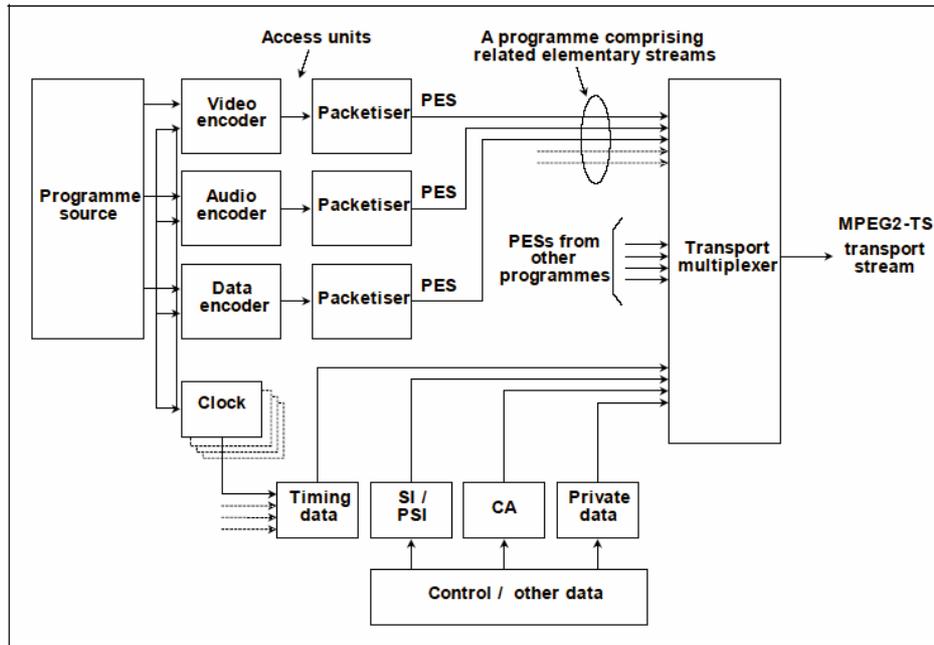


Figure 16 : Multiplexage MPEG2-TS

II.2.1.2 Paquet MPEG2-TS

Quelque soit la méthode d'encapsulation, le format de l'en-tête MPEG2-TS (au moins 4 octets) reste le même (cf. Figure 17). Les champs nous intéressant plus particulièrement sont :

- **PUSI** qui indique si le début d'une donnée est encapsulé dans le paquet (Pour le data streaming, cela correspond au début d'un PES)
- **PID** (13bits) : qui permet d'identifier un canal logique (flux élémentaire) au sein d'un multiplexe (Certains sont dédiés à la signalisation)
- **Adaptation Field** dans lequel on peut préciser la présence de données privées

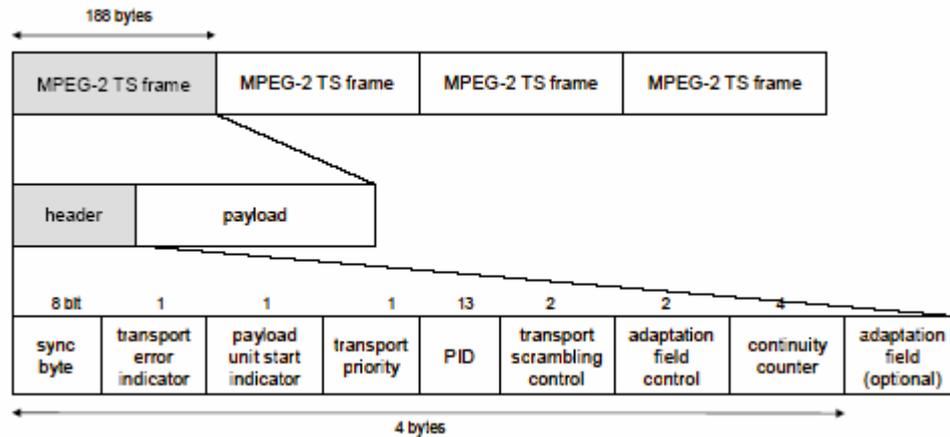


Figure 17 : Format d'un paquet MPEG2-TS

II.2.1.3 La signalisation DVB-S : les tables PSI/SI

Un récepteur DVB-S est ainsi à même de sélectionner un flux élémentaire au sein d'un multiplexe en fonction du PID associé à l'en-tête de chaque paquet MPEG2-TS. Un programme se constituant de plusieurs flux élémentaires, le récepteur doit être en mesure d'identifier les différents services disponibles : des PIDs liés à un même programme. Pour ce faire, le récepteur s'appuie sur les tables PSI/SI (Program and Service Information).

Nous nous contenterons dans cette partie de brièvement présenter les tables PSI/SI à travers un court exemple et recommanderons au lecteur de se référer aux normes suivantes pour de plus amples informations : la norme ISO MPEG2 [76] pour les tables obligatoires PSI, la norme DVB-SI [77] pour les tables optionnelles SI.

Afin d'illustrer l'emploi de ces tables, nous envisageons un exemple simple. Sur un satellite multi-transpondeurs, le terminal DVB-S de l'utilisateur se cale sur un canal de référence donnant l'accès à la table NIT (Network Information Table) qui regroupe l'ensemble des transpondeurs et des services rendus accessible à l'abonné par un opérateur. L'utilisateur choisit le multiplexe contenant les services l'intéressant. Le terminal se cale sur le transpondeur correspondant à partir des informations présentes dans la NIT. Le décodeur lit la table PAT (Program Association Table) qui lui permet de prendre connaissance de l'ensemble des services disponibles. A chaque service, est associée une table PMT (Program Map Table) qui référence les caractéristiques du service et de ses composantes. En « zappant » d'un programme à un autre, le terminal lit la PMT associée qui lui donne accès notamment aux PIDs des flux élémentaires audio et vidéo constituant ce programme.

Ces tables sont multiplexées au sein des flux de transport et sont émises pour la plupart périodiquement par l'entité de contrôle ou les fournisseurs d'accès. L'ETSI a défini une structure propre pour chaque table toute basée sur une structure commune : celle des sections DSM-CC (Digital Storage Media – Command and Control) [78]. Une section privée est transmise à travers plusieurs paquets MPEG2-TS utilisant un même canal logique au sein d'un multiplexe TS.

II.2.2 Méthode d'accès

Comme nous l'avons vu précédemment, les architectures traditionnelles DVB-S ne comprennent généralement qu'une unique GW n'émettant qu'un multiplexe occupant la largeur de bande totale du transpondeur visé. De ce fait, aucune méthode d'accès n'est spécifiée dans la norme DVB-S. Toutefois une GW peut transmettre plusieurs multiplexes. De même, un système DVB-S peut partager la bande passante d'un transpondeur entre plusieurs GWs. La méthode d'accès dans ce cas de figure est généralement propriétaire et s'appuie sur un multiplexage par répartition en fréquence ou TDM (Time Division Multiplexing).

II.2.3 Méthode d'encapsulation d'IP sur DVB-S

II.2.3.1 L'architecture d'encapsulation

La spécification MPEG2-TS offre plusieurs alternatives afin d'encapsuler des données privées (définies ni par le groupe DVB ni par le groupe MPEG2) autre que les trafics MPEG2 natifs, en particulier des paquets IP. On retrouve dans la Figure 18 l'ensemble des méthodes d'encapsulation possibles [75].

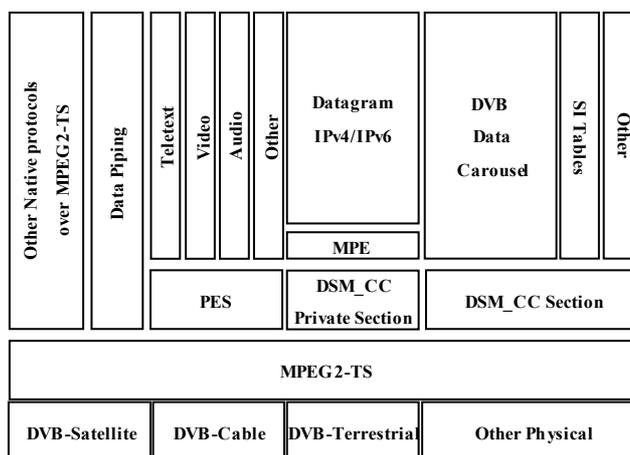


Figure 18 : Pile protocolaire DVB

Le mode *Data Streaming* que nous avons précédemment détaillé n'est pas véritablement adapté au transport de paquets IP qui ne forment généralement pas des flux continus. L'encapsulation de paquets IP dans des PES est néanmoins spécifiée, selon un mode asynchrone, par l'intermédiaire du champ PES *packet_data_byte_field* qui décline deux types de données privées (private stream 1 et 2).

Le mode *Data piping* offre un cadre pour les méthodes d'encapsulation propriétaire. Ainsi aucune méthode de fragmentation/réassemblage et encapsulation/désencapsulation n'est standardisée. Les nouveaux mécanismes d'encapsulation s'appuient généralement sur ce mode pour offrir des alternatives notamment à l'encapsulation d'IP dans MPE.

Le mode *Data Carousel* consiste en une émission cyclique de donnée vers des récepteurs. Ces données sont encapsulées périodiquement dans des sections privées DSM_CC de taille fixe appelé « section datagramme ».

Le mode *Multiple Protocol Encapsulation* (MPE) est le mécanisme d'encapsulation recommandé par la norme DVB-S pour le transport de datagrammes IP.

II.2.3.2 L'encapsulation MPE

L'encapsulation MPE fournit un mécanisme d'encapsulation des données au dessus de MPEG2-TS dans les réseaux DVB. Elle est conçue entre autres pour le transport de paquets IPv4 et IPv6. L'en-tête MPE contient un champ permettant de définir l'adresse du récepteur (adresse MAC) des données nécessaires quand de nombreux STs destinataires se partagent un même canal logique au sein d'un multiplexe MPEG2-TS. Les datagrammes IP sont encapsulés dans des sections DSM-CC compatibles avec le format de section privée MPEG2.

Si l'on considère le mode *section-packing*, les sections sont fragmentées et insérées à la suite afin de compléter les paquets MPEG2-TS. Aucun bit de bourrage n'est donc nécessaire. Afin de signifier la présence d'une nouvelle section dans un paquet MPEG2-TS, un *Pointer Field* (1 octet) est placé directement derrière l'en-tête MPEG2-TS. Il permet de préciser le nombre d'octets

entre le *Pointer Field* et le début de la nouvelle section utile.

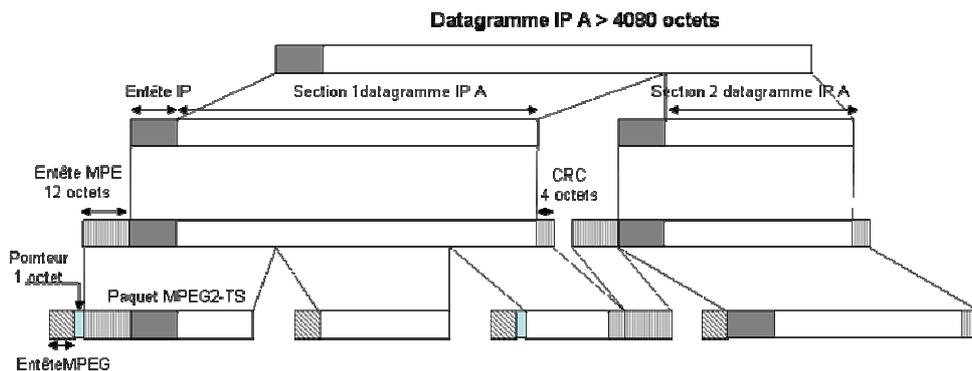


Figure 19 : Encapsulation d'un datagramme IP selon MPE

II.2.3.2.1 La signalisation

La norme [74] recommande l'utilisation de la table INT (IP/MAC Notification Table) afin de signaler la disponibilité et l'emplacement des flux IP au sein d'un multiplexe TS. Sous la notion de flux IP sont regroupés les séquences de paquets IP avec la même adresse IP source et/ou destination. A partir d'informations stipulées dans les tables NIT, PAT, PMT et INT, le récepteur va être en mesure d'évaluer le PID correspondant à un flux IP ou un ensemble de flux IP. A travers cette table, de nombreux descripteurs permettent de gérer les adresses IP dans un réseau DVB à travers notamment les identifiants suivants : *Platform_id*¹, *network_id*, *original_network_id*, *transport_stream_id*.

Cette norme reste toutefois relativement ouverte puisqu'elle se contente de définir une table d'association IP/PID mais ne propose pas de traitement automatique de cette dernière par le récepteur. Et le problème de la résolution d'adresse reste entier. Dans le cas des satellites, cette résolution d'adresse est un processus comprenant trois niveaux : l'adresse IP doit être résolue en une adresse MAC, ensuite doit être associée à un PID et enfin à un multiplexe TS spécifique. Quelque soit le niveau considéré, la table INT ne résout pas le problème de l'association des PIDs en fonction des adresses IP. Aucune association entre les adresses MAC et IP n'est spécifiée. Enfin la table INT ne peut être utilisé que dans le cadre d'une encapsulation MPE des paquets IP. Aujourd'hui cette méthode reste la seule implémentée dans les STs et l'association entre PIDs et adresses IP est généralement configurée statiquement.

II.2.3.3 L'alternative ULE (Unidirectional Lightweight Encapsulation)

Bien que l'encapsulation MPE soit recommandée par l'ETSI, elle présente des limites de performance. MPE n'a pas été uniquement conçu pour le transport d'IP. Aussi nombre de champs de l'en-tête MPE se révèlent inutiles. De ce fait, le groupe IP sur DVB de l'IETF s'est constitué notamment pour proposer une alternative à cette encapsulation : ULE [79]. ULE permet l'encapsulation directement sur MPEG2-TS de n'importe quel type de PDU via un SNDU (SubNetwork Data Unit) selon la méthode data piping défini par le groupe DVB (cf. Figure 20).

¹ la notion de plateforme IP/MAC regroupe des flux IP/MAC et/ou récepteurs et constitue ainsi un espace d'adressage sans conflit. Cette plateforme peut être présente sur plusieurs multiplexes et sur plusieurs réseaux DVB. Par ailleurs, plusieurs plateformes coexistent au sein d'un multiplexe TS.

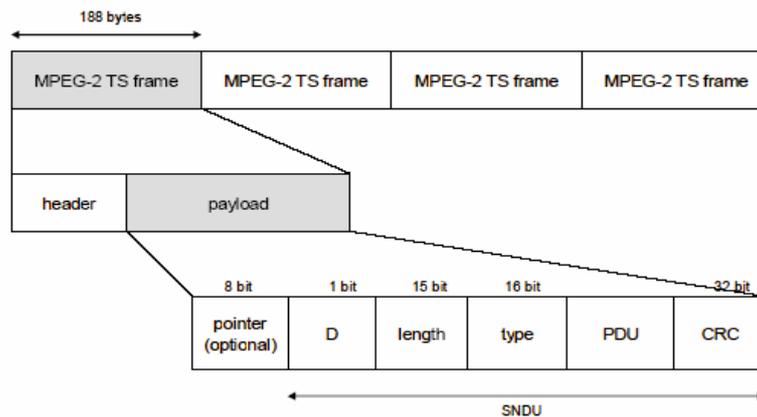


Figure 20 : Encapsulation ULE

L'en-tête SNDU est composé de différents champs dont :

- **indicateur D** indiquant la présence ou non d'un champ d'adresse destination MAC. Ce champ n'est qu'optionnel dans la mesure où le filtrage des paquets chez le récepteur peut uniquement se faire sur le PID et l'adresse IP destination.
- **taille** définissant la longueur du paquet
- **type** du PDU : IPv4 (0x0800), IPv6 (0x86DD), Ethernet ...
- **adresse destination** (optionnel) sera inséré entre le champ type et le PDU indique l'adresse MAC du récepteur du SNDU sur 48 bits

Afin de signifier la présence d'un nouveau SNDU dans un paquet MPEG2-TS, ULE n'a pas recours à l'Adaptation Field de MPEG2-TS (Adaptation Field Control à 01). Si un nouveau SNDU est contenu dans un paquet MPEG2-TS, le bit PUSI est mis à 1 et ULE définit un *pointer field* inséré derrière l'en-tête MPEG2-TS indiquant le bit de départ du nouveau SNDU.

Des études de performance ont déjà été effectuées démontrant la supériorité d'ULE sur MPE pour l'encapsulation d'IP [80][81]. Ce gain repose notamment sur la réduction des champs d'en-tête superflus mais elle présente également de nombreux avantages par rapport à MPE en ce qui concerne IPv6, la compression d'en-tête IP et l'IP Multicast.

Par ailleurs, le groupe IP over DVB propose également des mécanismes afin de lier les adresses IP au multiplexe TS [82]. Les solutions dynamiques s'appuient soit sur l'exploitation des tables MPEG2-TS INT ou MMT (Multicast Mapping Table) soit sur un véritable protocole de résolution d'adresse IP satellite. Cependant ces travaux ne sont pas encore assez matures pour être directement exploités.

II.3 La norme DVB-RCS

Avec la norme DVB-S, les terminaux sont uniquement en mesure de recevoir du trafic provenant du satellite. Au fur et à mesure que les services IP (Accès Internet large bande, Voix sur IP, Vidéo à la demande...) ont pris de l'envergure dans les réseaux terrestres, l'idée de supporter un canal interactif permettant au terminal utilisateur d'instaurer un véritable dialogue avec les serveurs et les utilisateurs d'Internet est apparue. La nature intrinsèquement asymétrique de ce dialogue et la nécessité d'un accès concurrentiel à la voie de retour afin d'éviter une monopolisation de la bande passante par une station unique exclut l'utilisation de la norme DVB-S qui ne définit aucune méthode d'accès partagé au médium et nécessite des équipements dont les prix prohibitifs ne peuvent être assumés par un utilisateur final. Ainsi, le standard UDLR (UniDirectional Link Routing) [83] a permis d'émuler à moindre coût une solution bidirectionnelle à travers un lien de retour terrestre, généralement un lien modem RTC classique. Cependant le réseau satellite alors lié à une infrastructure terrestre n'est plus complètement

autonome d'où la nécessité d'une voie de retour par satellite. La norme DVB-RCS (Digital Video Broadcast Return Channel over Satellite) spécifiée dans [67][84] standardise cette voie de retour via satellite pour les terminaux satellites. Couplée avec la voie aller s'appuyant sur la norme DVB-S, DVB-RCS introduit l'interactivité nécessaire à la définition d'une architecture d'un système multimédia large bande de nouvelle génération.

La norme DVB-RCS constitue la brique élémentaire des systèmes satellites bidirectionnels puisqu'elle permet au ST d'interagir. Comme nous l'avons vu précédemment, cette capacité d'interaction nécessite le support d'un nouveau modèle d'accès aux ressources sur la liaison montante.

II.3.1 Méthode d'accès : MF-TDMA

La méthode d'accès aux ressources de la voie retour est basée sur MF-TDMA (Multiple Frequency-Time Division Multiple Access). Ce système suppose une décomposition fréquentielle du lien de retour en plusieurs bandes de fréquence, elles-mêmes partagées en slots temporels durant lesquels les STs sont à même de transmettre des bursts de données.

II.3.1.1 *Les différents bursts RCS*

Il existe quatre types de burst DVB-RCS

- Des bursts de trafic basés soit sur des cellules ATM (ATM TRF burst) soit sur des paquets MPEG2-TS (MPEG2-TS TRF burst)
- le burst CSC permettant aux ST de s'identifier auprès du NCC lors de la phase de logon
- le burst ACQ utilisé si besoin lors de la procédure de synchronisation
- le burst SYNC utilisé pour maintenir la synchronisation ou envoyer des informations de contrôle vers le NCC

Un burst de cellules ATM est ainsi constitué d'une concaténation de cellules ATM (dont le nombre dépend de la taille de l'unité temporelle d'émission), d'un champ optionnel nommé SAC (Satellite Access Control) et d'un en-tête obligatoire.

Ce champ SAC se retrouve également dans le burst SYNC périodiquement envoyé durant la durée de la connexion du ST afin de le maintenir synchronisé.

Les autres bursts ACQ et CSC ont leur propre syntaxe d'encapsulation.

II.3.1.2 *Segmentation des ressources : Time-slots, Trames et supertrames*

La répartition des time-slots d'un canal MF-TDMA est centralisée par le NCC qui alloue périodiquement aux STs des séries de bursts, chacun défini par une fréquence, une bande passante, un temps de départ et une durée.

La Figure 21 illustre la composition d'une supertrame DVB-RCS. Les time-slots sont organisés en trames, elles-mêmes rassemblées au sein de supertrames. Une supertrame délimite l'ensemble des ressources d'un transpondeur DVB-RCS partagé par un groupe de STs.

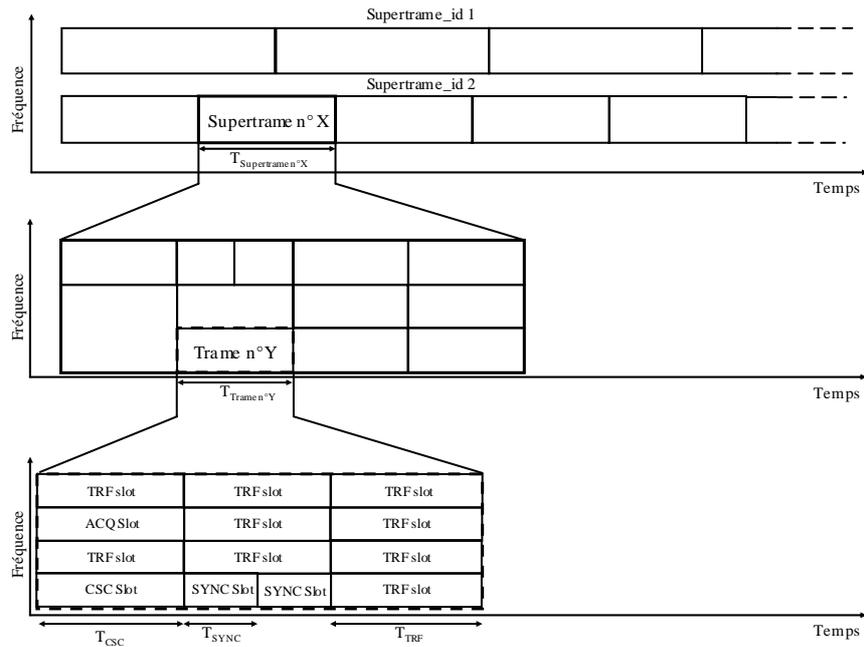


Figure 21 : Composition d'une supertrame DVB-RCS

A chaque supertrame est associé un plan d'allocation des time-slots la constituant entre STs concurrents. Les trames au sein d'une supertrame n'ont pas forcément la même durée, ni la même décomposition en time-slots.

Deux modes de partage MF-TDMA sont définis : l'un se basant sur des time-slots fixes (MF-TDMA statique) et l'autre sur des time-slots dynamiques (MF-TDMA dynamique). Dans le mode statique, la durée et la bande passante des time-slots successifs utilisés par un ST sont fixées. Une modification de leurs caractéristiques ne pourra être appliquée que dans une nouvelle supertrame. Dans le mode dynamique, qui n'est qu'optionnel, la bande passante et la durée des time-slots allouées à un ST peuvent varier au sein d'une supertrame. Ainsi le ST, en plus de s'adapter au changement de fréquence porteuse et de durée des time-slots, peut également être amené à modifier son débit de transmission et son type de codage d'un burst à un autre. Ce mode permet d'adapter au mieux les caractéristiques des time-slots aux divers profils de trafic multimédia ainsi qu'aux conditions de propagation. Cependant cette flexibilité accrue se fait au détriment des temps d'accès au support et demande des capacités de calcul supplémentaires de la part du NCC et des STs.

II.3.1.3 La pile protocolaire dans le plan de données

Les systèmes DVB-S/RCS sont basés pour la voie aller sur DVB-S et sur la voie retour sur DVB-RCS. Sur la voie aller, trois architectures protocolaires incluant IP sont autorisées :

- IP dans AAL5/ATM (selon la spécification [85]) dans un burst de trafic,
- IP dans AAL5/ATM dans MPEG2 TS (selon la spécification [86]) dans un burst de trafic,
- IP dans MPE dans MPEG2-TS dans un burst de trafic (méthode optionnelle définie dans [67])

Sur la voie retour, IP peut être encapsulé dans des bursts soit de type ATM, soit MPEG2-TS. L'architecture protocolaire illustrée dans la Figure 22 est celle adoptée si un satellite régénératif est utilisé.

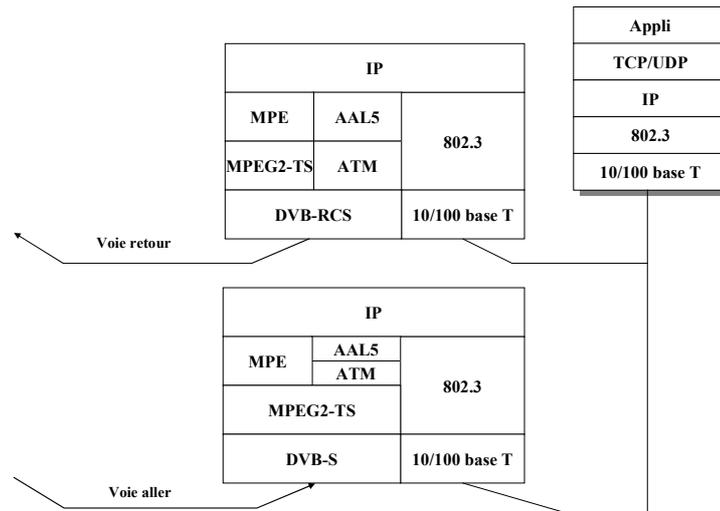


Figure 22 : Architecture protocolaire DVB-S/RCS dans le plan de donnée

II.3.2 La signalisation dans un système DVB-RCS/S

II.3.2.1 La signalisation sur la voie aller : des tables spécifiques DVB-RCS

Les tables SI véhiculées dans les flux de transport MPEG2-TS sur la voie aller contiennent l'ensemble des paramètres permettant aux récepteurs de démultiplexer et décoder les différents flux dans le multiplexe. Ces données sont structurées en tables encapsulées dans des sections privées DSM_CC, selon MPE, et véhiculées dans des canaux logiques aux PIDs prédéfinis. Les tables à destination des ST peuvent être distinguées selon 2 catégories :

- Les tables PSI/SI détaillées en section II.2.1.3
- Les nouvelles tables SI nécessaires au bon fonctionnement du protocole DVB-RCS : Superframe Composition Table (SCT), Time slot Composition Table (TCT), Satellite Position Table (SPT), Frame Composition Table (FCT), Terminal Burst Time Plan (TBTP), Correction Message Table (CMT)

Ainsi le SCT décrit le partitionnement complet des ressources satellite en supertrames et le positionnement relatif des trames les unes par rapport aux autres au sein des supertrames. La table FCT contient pour chaque trame la décomposition en time-slots et leur emplacement respectif. Chacun de ces time-slots est décrit par la TCT qui contient toutes les caractéristiques techniques du time-slot et le type de burst contenu (ATM ou MPEG2-TS TRF, SYNC, ACQ, CSC). Enfin la table TBTP contient le plan d'allocation des time-slots d'une supertrame partagée entre un groupe de STs. La table SPT indique la position du satellite et la table CMT permet aux STs d'effectuer les corrections nécessaires pour leurs prochaines émissions.

Les messages TIM (Terminal Information Message) sont des messages privés envoyés par le NCC à destination d'un ST ou d'un groupe de STs. Ils servent notamment à transporter des identifiants logiques (Group_id, Logon_id, Channels_ids, ST IP addresses, PID values, PVCs), des messages d'échec ou de réussite de la procédure de logon.

Enfin la NCR (Network Clock Reference) permet aux STs de se synchroniser.

L'ensemble de ces trames est répertorié dans la Figure 23.

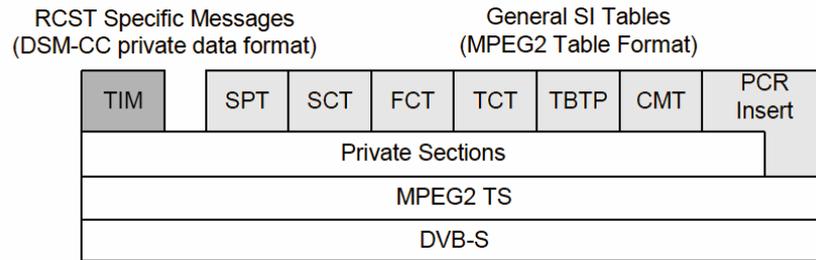


Figure 23 : Pile protocolaire pour la signalisation RCS sur la voie aller

II.3.2.2 La signalisation sur la voie retour

II.3.2.2.1 Les procédures de logon et de synchronisation

A son démarrage, le ST entame une procédure de logon afin d'accéder aux ressources du lien satellite. Le terminal satellite émet à destination du NCC une première requête d'accès par un burst CSC. Ce burst inclut l'adresse MAC du ST et des informations concernant ses capacités. Après l'acceptation du NCC, ce dernier répond par un message TIM contenant les identifiants logiques associés au ST : VPI (Virtual Path Identifier) et VCI (Virtual Connection Identifier) ATM et PIDs pour émettre du trafic et échanger les messages de contrôle et d'administration (MPEG2-TS burst). Le NCC lui attribue également un time-slot dédié à la signalisation de synchronisation qui notamment va permettre au terminal de se maintenir synchronisé sur l'horloge de référence NCR par l'intermédiaire de message périodique SYNC.

Les informations échangées pendant la phase de logon peuvent être utilisées afin de nourrir une base de donnée dans le NCC et le ST permettant de simplifier la résolution d'adresse. L'attribution d'une adresse IP et de l'adresse IP de la GW par défaut n'est pas spécifiée par la norme et donc repose généralement sur une configuration manuelle ou par SNMP.

II.3.2.2.2 Les catégories de requêtes de capacité

L'un des principaux intérêts de la norme DVB-RCS est d'offrir un canal retour dont les capacités peuvent varier au cours du temps. Chaque ST est en mesure de requérir dynamiquement des « capacités de transmission » auprès du NCC qui les lui alloue en fonction des besoins exprimés et de l'état des ressources disponibles. Ainsi si, à un instant donné, une accumulation de trafic à émettre est observée dans un ST, celui-ci peut requérir plus de bande passante auprès du NCC. Cet algorithme est le DAMA (Demand Assigned Multiple Access). Il dispose ainsi de 5 catégories de capacité dont trois 3 types de requête de capacité (CR, Capacity Request):

- Continuous Rate Assignment (CRA) : une quantité de time-slots fixe par supertrame négociée au logon du ST et allouée pour la durée de connexion du ST.
- Rate Based Dynamic Capacity (RBDC) : Une quantité de time-slots négociée pour une supertrame ne pouvant excéder un seuil max ($RBDC_{max}$). Bien qu'optionnelle, cette capacité permet de compléter ponctuellement un minimum de capacité alloué par le CRA
- Volume Based Dynamic Capacity (VBDC) : Une quantité de time-slots négociée auprès du NCC et qui peut être répartie sur plusieurs supertrames. Ces requêtes sont cumulatives.
- Absolute Volume Based Dynamic Capacity (AVBDC) : Une quantité de time-slots négociée auprès du NCC qui peut être répartie sur plusieurs supertrames. Une nouvelle requête AVBDC annule la précédente.
- Free Capacity Assignment (FCA): ce type de capacité représente les time-slots d'une supertrame restant après le traitement des autres types de capacités par le NCC et qui sont distribués équitablement ou non entre les STs concurrents à hauteur d'un seuil prédéfini.

L'attribution des time-slots se fait à chaque supertrame et respecte l'échelle suivante de priorité

entre les types de capacités : CRA>RBDC>A(VBDC)>FCA.

Afin de transporter les CRs vers le NCC, deux types de signalisation sont supportés : une signalisation intra-bande et une signalisation hors-bande. La signalisation intra-bande s'appuie sur une méthode liée à l'utilisation du champ optionnel SAC (Satellite Access Control) attaché au burst de trafic ou sur la méthode DULM (Data Unit Labelling Method) qui permet au ST de transmettre des informations de contrôle et/ou d'administration au NCC dans des bursts normalement dédiés au trafic. La signalisation hors bande s'appuie elle sur la méthode « mini-slot » avec ou sans contention qui correspond à l'allocation périodique à un ST ou un groupe de STs de bursts de plus courte durée que les bursts de trafic.

Cette allocation de bande passant à la demande constitue incontestablement un des points forts des systèmes satellites en terme de QoS en répartissant la bande passante totale d'un réseau satellite entre les STs en fonction de leur charge ponctuelle comme nous le verrons dans la section II.4.3.

II.3.3 Intelligence embarquée

II.3.3.1 *Commutation à bord*

Comme nous l'avons vu dans la section II.1.3, le support de services bidirectionnels multimédias large bande dans les réseaux satellite d'accès de nouvelle génération sera grandement facilité par l'intégration dans le satellite de techniques multifaisceaux et d'une charge utile régénérative. La combinaison de ces deux techniques permet d'économiser grandement les ressources du satellite.

D'une part, les techniques multifaisceaux en divisant ainsi la couverture globale du satellite en de multiples faisceaux restreints permettent un gain considérable en terme de bande passante par :

- la réutilisation des fréquences d'un spot à un autre
- la circonscription du trafic unicast à un spot restreint contrairement à l'inondation superflue de nombreux terminaux dans le cadre d'un satellite monofaisceau
- une réduction des coûts des terminaux et de l'envergure de leur antenne

D'autre part, la charge utile intégrée dans le satellite va permettre une gestion flexible et dynamique des connexions entre les différents spots offrant ainsi une gestion fine et dynamique des ressources basée sur les connexions et des communications inter-STs en un simple bond.

Mais qu'entend-on véritablement par charge utile régénérative ou intelligence embarquée ?

L'intelligence embarquée consiste au déport de quelques fonctionnalités de contrôle et d'administration habituellement effectuée au sol au sein du satellite. L'avantage évident de ce déport est une réduction de moitié des délais de propagation de la signalisation et de la charge de signalisation. Cependant le problème est que chaque fonctionnalité déportée introduit un coût, un poids et une complexité supplémentaires.

Les systèmes tels IBIS [87] ou AmerHis [88] s'appuient ainsi sur des satellites embarquant des commutateurs de type circuit et regroupant :

- des multiplexeurs/démultiplexeurs permettant de régénérer les paquets issus des voies montantes et de les remultiplexer dans les flux de transport descendants correspondants
- un commutateur doté d'autant d'entrées et de sorties que de spots ascendants et descendants
- un bloc de contrôle permettant de configurer l'OBP et comprenant notamment la table de commutation

La Figure 24 illustre l'ensemble des piles protocolaires possibles dans ce type de configuration régénérative. Les systèmes embarqués envisagés dans le projet IBIS et implémentés dans AmerHis se basent sur un commutateur MPEG2-TS de circuits. Ce type de commutateur s'appuie ainsi sur un multiplexage par position. Le multiplexage par position implique une

commutation des time-slots basée sur leur position temporelle et fréquentielle respective dans la supertrame. Les en-têtes ne sont pas utilisés dans ce processus. La table de multiplexage est préconfigurée et lie directement les slots du multiplexe montant aux slots du multiplexe descendant.

Architecture protocolaire

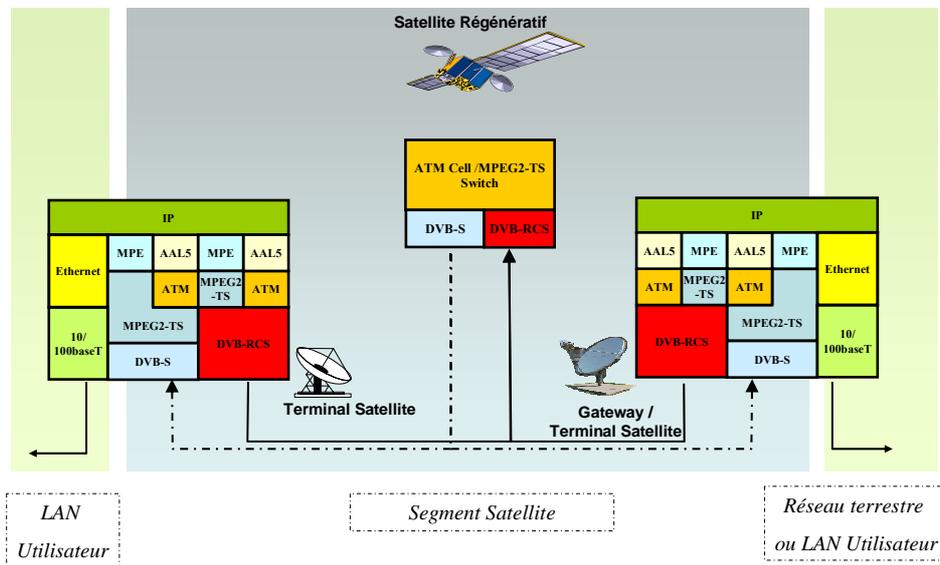


Figure 24 : Architecture protocolaire d'un système régénératif

La Figure 25 donne un aperçu de la structure de la charge utile spécifiée dans le projet IBIS. Chaque transpondeur DVB-RCS est implémenté comme un Multi-Carrier Demultiplexer, Demodulator and Decoder (MCDDD) permettant de restituer les différents paquets MPEG2-TS issus des différentes supertrames constituant un multiplexe MF-TDMA. Ces paquets sont ensuite orientés vers les multiplexes descendants correspondants.

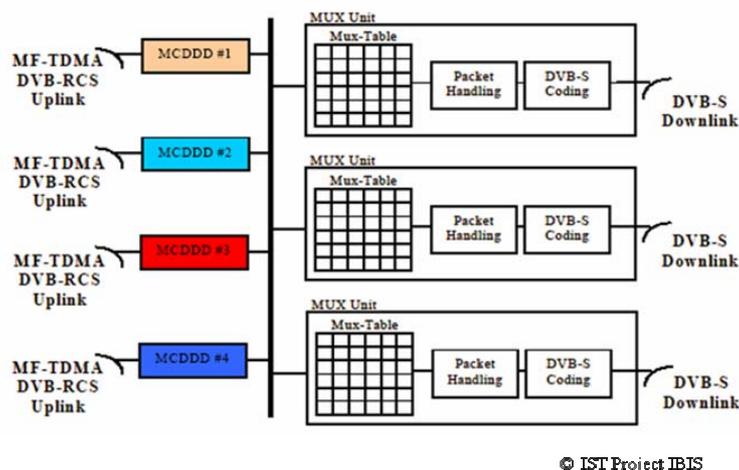


Figure 25 : Charge utile régénérative

Hormis les données à acheminer d'un spot à un autre, il nous faut également considérer la signalisation au logon et durant la connexion du ST (DAMA notamment). Comme nous l'avons

souligné dans la section II.1.3.1, le NCC et le NMS sont découplés de la GW dans ces configurations régénératives et sont regroupés au sein d'un unique équipement qui peut être situé dans son propre spot, le « spot service » où la signalisation du plan de contrôle et d'administration (CTRL/MNGM) converge. Nous l'opposons au « spot utilisateur » destiné aux données du plan utilisateur. Nous noterons toutefois que le spot service est très souvent confondu avec un « spot utilisateur » ce qui permet d'économiser une antenne satellite.

En dehors de la commutation de circuit, la commutation de paquet est également envisagée dans des architectures telles que DIPCAST [89] qui s'appuie sur un brasseur ATM. Contrairement à la commutation de circuit, le brasseur DIPCAST implémente un multiplexage par étiquette. Le brasseur ATM implémenté commute les cellules ATM de la voie montante en fonction de leur champ VPI et d'une table de commutation mise à jour à l'ouverture et à la fermeture de connexions si besoin est.

Deux choix sont généralement conseillés pour assurer cette commutation par paquet :

- un mode *self-switching* où les labels de destination associés aux cellules ATM ou paquets MPEG2-TS sont construits selon un plan d'adressage hiérarchique avec un certain nombre de bits identifiant les spots descendants ou les ports de sortie du commutateur et le reste constituant des identificateurs de terminaux ou de groupes de terminaux.
- un mode *label-switching* utilisant des labels et une table de commutation embarquée. Afin de commuter les cellules ATM ou les paquets MPEG2-TS, seule une partie du champ dédié au PID ou VPI/VCI est nécessaire à l'instar du brassage ATM qui se base uniquement sur le champ VPI. Les informations de niveau connexion sont à proscrire pour effectuer la commutation afin de réduire la quantité d'informations à traiter par cellule ou paquet et la fréquence de mise à jour des tables.

Enfin, au niveau des STs récepteurs, les champs VCI/VPI et PID/MAC sont utilisés pour déterminer l'adresse du ST source ou d'un groupe de STs de destination.

II.3.3.2 L'adressage

Quelque soit le type de commutation considérée, se posent les problèmes :

- du faible lot de labels disponibles
- de la mise à jour des caches d'association labels/adresses IP dans les STs
- de la mise à jour de la table de commutation dans l'OBP
- de la mise à jour de l'état des connexions dans le NCC

D'ores et déjà, nous pouvons écarter une commutation à bord basée sur l'adresse destination MAC car cela supposerait la restitution de la section MPE ou du SNDU ULE, ces derniers étant fragmentés en paquets MPEG2-TS au niveau du ST.

Dans le cas de bursts TRF MPEG2-TS, l'attribution d'un PID pour chaque ST potentiel n'est pas envisageable dans la mesure où les 13 bits du PID ne donnent accès qu'à 8192 valeurs alors qu'on envisage jusqu'à 10000 STs simultanément connectés sur un même système satellite et qu'un nombre conséquent de PIDs sont dédiés à la signalisation. Ainsi si l'on se réfère au projet IBIS, il est question pour un futur système satellite de taille moyenne de supporter 10000 STs répartis en groupe de 500 dans 20 spots différents.

Avec des bursts TRF ATM, les champs VPI (12bits) et VCI (16bits) d'une cellule type NNI (ou UNI en considérant le champ GFC) offrent des possibilités accrues comme nous le verrons dans l'architecture SATIP6 [90] détaillée dans le Chapitre 3.

La gestion de ces labels doit toutefois se faire avec parcimonie si l'on veut pouvoir proposer une solution supportant la mise à l'échelle aussi bien en terme de labels disponibles qu'en terme de signalisation (mise à jour des différents caches, de la table de commutation de l'OBP et des ressources attribuées par connexion au niveau du NCC). Les enjeux à prendre en compte sont multiples :

- Le nombre de terminaux et de spots montants et descendants supportés

- L'ensemble des types de connexions possibles dans cet environnement (connexions point-à-point, point-à-multipoint et multipoint-à-multipoint)
- Le support de différentes classes de service (CdS) au niveau 2
- La résolution d'adresse au niveau émetteur pour connaître l'adresse MAC destinatrice et le PID ou le VPI/VCI à utiliser pour émettre. De même, elle servira au niveau du récepteur afin d'effectuer un filtrage à plusieurs niveaux (PID/MAC ou VPI/VCI et IP)

Les champs PID ou VPI/VCI contiennent ainsi des labels identifiant des réseaux virtuels de niveau 2. Un label peut correspondre à un groupe de STs (VPN, une partie d'un VPN ou un groupe multicast), à une GW pour l'accès à Internet ou bien même un terminal satellite. Ainsi, selon le niveau d'association envisagé, le ST sera en mesure ou non de filtrer les paquets ou cellules au niveau 2. Dans un mode connecté, le ST sera à même de différencier les connexions tandis que dans un mode non connecté le ST sera obligé de reconstituer les paquets IP avant de pouvoir déterminer si ces cellules ATM et paquets MPEG2-TS lui sont destinés.

Enfin une distribution intelligente des adresses IP est également envisageable si elle est gérée par un même INAP¹ (Interactive Network Access Provider) afin d'offrir, par exemple, des plages d'adresses IP contiguës pour des entreprises localisées sur différents spots et reliées par un même VPN.

En ce qui concerne les mécanismes de résolution d'adresse, il faut notamment compter sur une table de correspondance au sein de chaque ST mis à jour par deux mécanismes :

- Une table de signalisation MPEG2-TS transmise au niveau 2 (MMT pour le multicast et INT pour l'unicast)
- Un protocole de résolution d'adresse transporté par IP

II.3.3.3 *Gestion des ressources sur la voie montante*

Dans un souci de clarté, nous considérons que le NCC ne gère les ressources que d'un unique IN (Interactive Network)² et que les STs associés à cet IN ne peuvent s'abonner qu'à un unique SP. Nous supposons également que les liaisons descendantes sont dimensionnées afin de supporter l'ensemble du trafic provenant des liaisons montantes.

Tout d'abord les ressources de chaque multiplexe MF-TDMA sont gérées indépendamment au sein du NCC.

Ensuite, pour chaque multiplexe MF-TDMA, nous distinguerons les ressources associées à un *canal MAC*. Cette notion bien que définie dans la norme DVB-RCS est particulièrement vague et ne permet pas de faire le lien direct entre un canal et les ressources satellites. C'est pour cela que nous reprenons la définition du projet AmerHis [91] en l'étendant. L'ensemble des time-slots d'un multiplexe MF-TDMA commutés vers un ou plusieurs spots descendants constitue un *canal MAC*. Ce dernier est ainsi défini par une source (un spot montant) et une ou plusieurs destinations (un ou plusieurs spots descendants). Ainsi, dans un système tel qu'AmerHis, la commutation de circuit implique que les time-slots associées à un canal au sein d'un multiplexe MF-TDMA lui sont dédiées et ne peuvent être partagées avec d'autres canaux. Par contre, dans le cadre d'une commutation par paquet, les time-slots alloués à un canal durant une supertrame peuvent être partagés avec d'autres canaux lors de la prochaine supertrame.

A l'instar d'un conduit logique ATM (Virtual Path), le *canal MAC* regroupe plusieurs connexions ou voies logiques ATM (Virtual Connection) ou canaux logiques MPEG2-TS. Ces connexions sont définies par un ST source et un ou plusieurs STs destination.

Ainsi, dans la Figure 26, nous observons ces différentes notions dans le cadre d'AmerHis (commutation par circuit, commutateur MPEG2-TS). Un canal allant d'un ST_A au TDM₁ (Time

¹ Un INAP est l'opérateur d'un IN (Interactive Network)

² Un Interactive Network représente une partie des ressources d'un réseau satellite partitionné par l'opérateur du réseau satellite (SNO) affecté à un INAP (Interactive Network Access Provider). L'INAP est l'entité avec laquelle les SPs (fournisseur de service) négocient l'accès au réseau satellite.

Division Multiplex) est identifié en tant que Channel_ID-1. Au sein de ce canal, deux classes de service, HP (High Priority) et LP (Low Priority), sont différenciées par la valeur du PID. Enfin les connexions du ST_A vers des STs (ST1, ST2 et ST3) du TDM₁ sont identifiées par le PID et l'adresse de destination MAC. Dans un contexte ATM, ces connexions se concrétisent par des PVCs (ATM Permanent Virtual Channel) renouvelées en fonction de l'activation de tel ou tel service applicatif.

Enfin les flux IP provenant d'un LAN externe relié au ST sont transportés à travers ces connexions jusqu'à un autre ST.

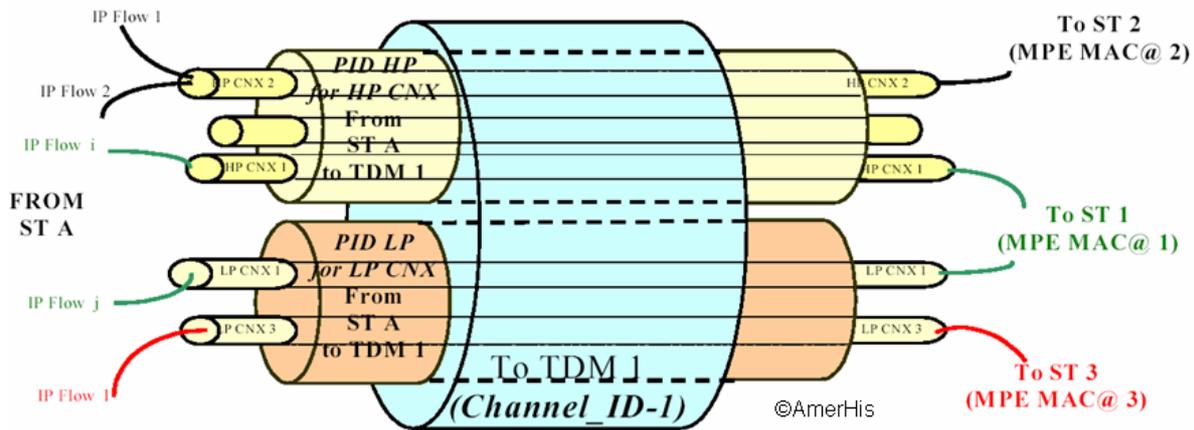


Figure 26 : Segmentation hiérarchique des ressources d'une voie montante

Cette segmentation est toujours valable dans un contexte non régénératif. La seule différence est que les STs n'ont accès qu'à un seul canal dont l'unique destination possible est la GW associée et que la GW intègre alors les fonctions de NCC et NMC.

II.3.4 Les connexions

Dans l'Annexe J de [84], la notion de connexion et le protocole de contrôle, Connection Control Protocol (C2P), sont développés. Outre les différents types de connexions possibles (uni- et bi-directionnel, point- ou multipoint-à-unipoint ou -multipoint), les différents identifiants sont détaillés (Channel_id, Source & Destination Address, Forward and Return Stream Identifiant).

Dans le projet AmerHis, trois types de connexions sont distingués afin de répondre aux besoins de différents types de services et à différents modes de gestion des ressources des INs:

- Les connexions permanentes établies au logon pour la durée de la connexion du ST avec des caractéristiques de base (CRA, RBDC_{max}, ...) définies par l'intermédiaire d'un SLA qui peuvent être renégociés au cours de la connexion du STs
- Les connexions semi-permanentes négociées au logon du ST mais dont l'activation est différée dans le temps pour répondre aux besoins de services périodiques
- Les connexions à la demande impliquant une requête dynamique de la part du ST nécessitant un contrôle d'admission et une réservation de ressource auprès du NCC. Ces connexions sont utilisées pour des services ponctuels.

C2P peut ainsi être utilisé afin de modifier les paramètres d'une connexion permanente ou pour requérir une connexion à la demande. La norme DVB-RCS est loin d'avoir complètement spécifié C2P. Actuellement les requêtes de ST à NCC sont codées dans des champs IE (Information Elements) véhiculés au format DULM et permettent d'établir/modifier/fermer une connexion ou un canal. Le NCC quant à lui répond à travers différents descripteurs des messages TIM. Cependant des alternatives à cette signalisation de niveau 2 sont considérées à travers des

protocoles de plus haut niveau.

Afin de mieux situer notre contribution, nous allons pour finir recenser les notions de QoS liées aux systèmes satellites.

II.4 Architecture de QoS dans les réseaux d'accès satellite DVB-S/RCS

Les réseaux satellites appartiennent aujourd'hui aux technologies d'accès au même titre que l'ADSL, l'UMTS ou le Wimax. La baisse du coût des terminaux, les efforts soutenus en matière de standardisation et d'interopérabilité, l'évolution des techniques de codage et de compression, tout semble converger pour doter le satellite des moyens nécessaires à sa future intégration.

Cependant, d'un point de vue architecture réseau et notamment pile protocolaire, les réseaux satellites ne sont pas encore prêts à être intégrés à l'infrastructure NGN. En effet, à l'instar des autres technologies d'accès, les protocoles IP largement éprouvés sur les réseaux terrestres dédiés à la signalisation, à la gestion et au contrôle nécessitent d'être adaptés aux spécificités des réseaux satellites qui souffrent notamment d'un produit délai-bande passante large particulièrement dommageable à l'efficacité de TCP [92]. Ainsi beaucoup de protocoles de gestion de ressource et de signalisation sont liés à TCP (COPS, SIP) et des solutions tels que les Performance Enhancing Proxy [93] doivent être mis en place pour « adapter » TCP sur le segment satellite. Par ailleurs, pour la signalisation, des adaptations peuvent se révéler nécessaires pour sauvegarder les ressources du satellite : compression des messages de signalisation, agrégation des réservations de ressource, adaptation des timers de retransmission.

Enfin, les protocoles de signalisation de niveau IP liés à la QoS (politique de QoS, contrôle d'admission, provisionnement de SLAs) se doivent d'interagir avec les mécanismes de niveau 2 afin d'offrir une gestion optimale des ressources trop grossière au niveau IP.

C'est à partir de ces constatations que le groupe BSM de l'ETSI a décidé de définir un cadre architectural propice à l'intégration de ces protocoles adaptés aux spécificités des réseaux BSM [94]. Ce modèle architectural se veut générique et non spécifique à un système satellite particulier (Type DVB-RCS). Le groupe BSM a ainsi clairement identifié une entité « Protocol Manager » dont le rôle est d'intercepter ces protocoles extérieurs et de vérifier s'ils nécessitent d'être adaptés¹ sur le segment satellite. Par ailleurs un modèle d'acteurs a également été détaillé (cf. II.4.1) afin de mieux cerner les interactions entre les fournisseurs de service, les consommateurs et les différents opérateurs du réseau satellite.

L'ensemble des recommandations sur lesquelles nous nous appuyerons dans cette section sont issues de la norme DVB-RCS et des travaux menés par le groupe de travail BSM de l'ETSI.

II.4.1 Les différents acteurs du réseau satellite

Un des éléments essentiels de l'intégration d'un réseau dans les NGNs est de bien séparer les différents acteurs concourant à la fourniture d'un service dans ce réseau. Pour les réseaux satellites appartenant aux réseaux de télécommunication traditionnels, ce processus se révèle d'autant plus important qu'ils sont encore largement tributaires de leur forte intégration verticale passée. Ces acteurs ont ainsi été clairement identifiés dans le rapport technique du BSM ETSI TR 101 984 [95] sur les services et les architectures de référence pour les systèmes BSM :

- L'opérateur du réseau satellite (SNO) : il possède et gère la maintenance, la gestion et le déploiement du système DVB-RCS à l'exception des terminaux (STs & GWs). Il gère également le partitionnement des ressources entre les INAPs en fonction de leur contrat
- L'opérateur du satellite : Il est responsable du satellite et coopère, pour un système régénératif, avec le SNO pour la configuration de l'OBP.
- L'opérateur du réseau d'accès interactif (INAP) : Il a accès à une partie des ressources du satellite par l'intermédiaire d'un IN (Interactive Network) qu'il négocie auprès du SNO. Il

¹ Par adaptation, le groupe BSM entend le processus visant à améliorer les performances de ces protocoles sur le segment satellite en réduisant par exemple le nombre de messages transmis sur le segment satellite.

partage ensuite ces ressources entre différents SPs (Fournisseurs de Service) à travers des VSNs (Réseaux Satellite Virtuels). Un VSN représente les ressources offertes par un INAP à un SP. Le choix incombe toutefois au SP de gérer ou non lui-même ces ressources satellites. On distinguera ainsi deux types de contrats dénommés respectivement « Bandwidth » où les ressources attribuées sont dédiées au SP et « Wholesale Connectivity » où la gestion des ressources revient à l'INAP qui partage alors ces ressources entre plusieurs SPs.

- Les fournisseurs de service (SP) qui peuvent être les ISPs (Internet Service Provider) mais également des fournisseurs de services applicatifs ou ASPs (Application Service Provider), MSP (Fournisseurs de service Multicast), VPN SP (Fournisseurs de service VPN), ITSP (Fournisseurs de service téléphonique par Internet)...
- Les abonnés : Ils constituent l'entité intermédiaire à laquelle les utilisateurs peuvent déléguer le choix des fournisseurs de service. Il peut utiliser les services mis à disposition par l'INAP, et contracter plusieurs abonnements auprès de différents SPs afin de répondre aux exigences des utilisateurs
- Les utilisateurs : ils peuvent être directement reliés au réseau satellite par le ST ou par un LAN connecté à ce ST. Les terminaux utilisateurs sont connectés aux applications fournies par les différents SPs. L'ensemble du LAN et des terminaux utilisateurs sont référencés sous le nom de Customer Premises Equipment (CPE).

II.4.2 Le modèle architecturale BSM basé sur IP

L'architecture proposée dans la Figure 27 a été définie dans le rapport TR 101 984 [95]. Elle définit ainsi une interface SI-SAP (Satellite Independent Service Access Point) entre les fonctionnalités indépendantes du système satellite sous-jacent et les fonctionnalités spécifiques au système satellite utilisé. Deux fonctions d'adaptations s'ajoutent : l'une SIAF (Satellite Independent Adaptation Function) en bas de la couche 3 pour adapter les protocoles de niveau 3 aux « services support BSM » et inversement, et l'autre SDAF (Satellite Dependent Adaptation Function) qui opère en haut de la couche 2 pour adapter les « services support BSM » avec les services offerts par l'interface air native.

Ces services support BSM s'appuyant sur les mécanismes de niveau 2 du système satellite sous-jacent offrent notamment des services orientés ou non connexion, uni- ou bi-directionnel et symétrique ou asymétrique (en bande passante, en QoS...) et des connexions unicast ou multicast.

Les couches SLC et SMAC peuvent être apparentées au niveau de la couche 2 du modèle OSI et la couche SPHY à celui de la couche 1.

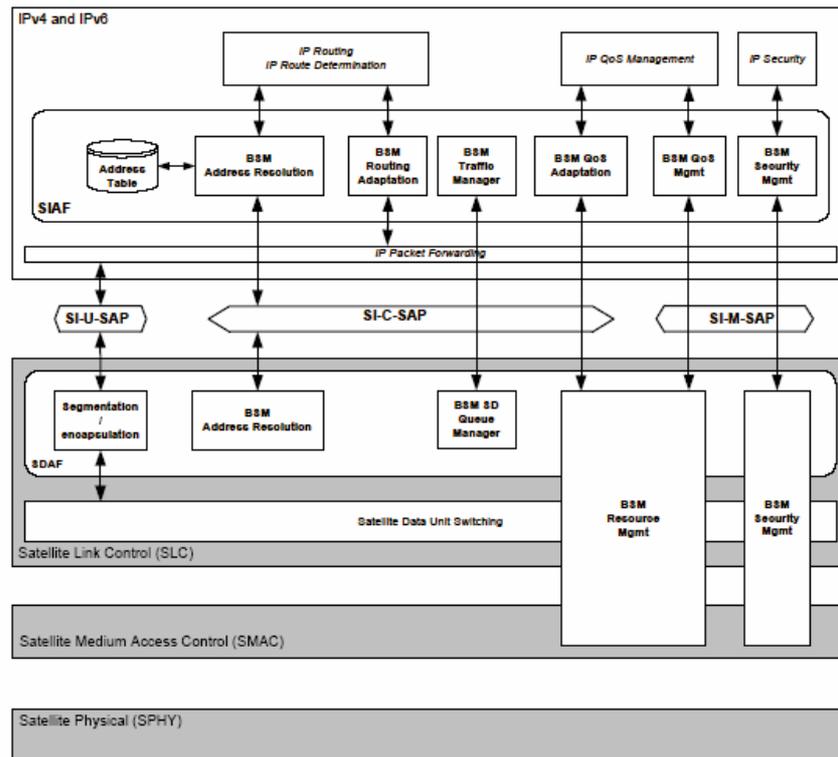


Figure 27 : Architecture protocolaire BSM pour service unicast

Par ailleurs, cette architecture définit également des interactions à différents niveaux avec les réseaux IP externes, notamment à travers les fonctionnalités de pont (niveau 2), de niveau IP (Routage) et de Gateway (pour les niveaux supérieurs à IP, les Performance Enhancing Proxies¹ en font parti). Les terminaux satellites à l'entrée et à la sortie du réseau satellite réalisent ainsi cette interaction entre les différents protocoles externes et internes.

Les normes DVB-S et -RCS définissent des mécanismes de niveau 1 et 2 que l'on peut assimiler aux couches SLC, SMAC et SPHY définis dans l'architecture BSM. Ainsi les systèmes satellites DVB-S/RCS constituent la première famille de systèmes satellites qui offre un cadre concret à la réalisation des concepts abstraits BSM.

Cependant, en ce qui concerne la QoS, le rapport TR 102 157 [96] se contente d'un recensement des mécanismes de QoS existants dans le terrestre (SIP, COPS, RSVP, DiffServ, Classes de Trafic, DAMA ...) mais ne préconise rien de nouveau concernant l'adaptation de ces mécanismes au contexte satellite. Elle offre toutefois un cadre architectural nécessaire à l'établissement de solutions génériques.

Détaillée dans les chapitres 3 et 4, la proposition, l'émulation et l'évaluation d'une architecture de QoS pour systèmes DVB-RCS/S intégrant les 3 niveaux MAC, IP et Application de manière cohérente constituera ainsi notre première contribution.

II.4.3 L'algorithme de DAMA : allocation de bande passante à la demande

Le principal mécanisme de niveau 2 avec lequel une architecture de QoS IP satellite va devoir interagir est le DAMA qui offre toute la souplesse nécessaire pour partager au mieux les ressources réduites disponibles sur la voie de retour en fonction des besoins ponctuels de chaque terminal. Jusqu'à présent, les opérateurs satellites actuels se contentent généralement de SLAs très simples fournissant à chaque ISP des ressources fixes entre terminaux (topologie maillée) et entre

¹ Les PEPs sont généralement utilisés dans le but d'améliorer les performances d'une connexion TCP sur un segment réseau aux caractéristiques particulières comme un segment satellite.

la GW et les terminaux (topologie d'accès) lui appartenant. Cette allocation statique des ressources a lieu à la connexion du terminal et elle reste inchangée pour la durée de la connexion. Ce provisionnement statique des ressources a le mérite d'être simple pour l'opérateur satellite mais entraîne un gaspillage dû à l'impossibilité de redistribuer les time-slots non consommés par un ST aux autres STs.

La norme DVB-RCS a ainsi spécifié un DAMA qui est un mécanisme d'allocation de ressources à la demande. Il fonctionne sur le principe client-serveur où le ST est doté du client DAMA et le NCC du serveur DAMA. Le NCC centralise ainsi la gestion de l'ensemble des ressources du réseau satellite et les alloue en fonction des demandes des différents STs. Comme nous l'avons vu dans la section II.3.2.2.2, le ST dispose de 5 catégories de capacités. La norme définit également des classes de trafic au niveau MAC qui peuvent être matérialisées ou non dans le ST par des files MAC distinctes. Ces classes sont au nombre de trois : la classe RT (Real Time) pour les applications à fortes contraintes temporelles, la classe VR (Variable Rate) décomposée en deux catégories VR-RT et VR-JT dédiés respectivement au trafic à débit variable sensible ou tolérant à la gigue et Jitter Tolerant (JT) pour le reste du trafic. Des types de capacités sont recommandées pour chaque classe : CRA pour RT, RBDC pour VR et VBDC/AVBDC pour JT.

L'intérêt de ces mécanismes est clair puisqu'il offre un moyen de différencier au niveau 2 des applications aux exigences en terme de QoS différentes et d'y répondre spécifiquement par classes de trafic. Il faut ainsi savoir que la latence du cycle de demande/allocation des ressources diffère d'une catégorie de CR à l'autre. Cette souplesse du DAMA permet ainsi de répondre à la fois aux garanties strictes de délai et de gigue nécessaires aux applications temps réel et d'économiser les ressources du satellite en partageant le reste des ressources de manière optimale entre les différents STs en fonction de leur besoin et de l'état de charge du réseau satellite.

Enfin la norme aborde succinctement le problème de l'allocation des time-slots accordés périodiquement par le NCC à travers le TBTP au sein d'un terminal.

Cependant le caractère particulièrement ouvert de ces recommandations laisse toutefois une place trop grande à l'interprétation et aux solutions propriétaires empêchant l'interopérabilité de tels systèmes et par la même l'utilisation systématique d'une gestion dynamique des ressources.

Des travaux sont actuellement en cours dans le groupe Satlabs de l'ESA (European Space Agency) [97] pour spécifier sur quel types d'informations les calculs des CRs doivent se baser et fournir des algorithmes en fonction des catégories de requête considérées comme base de référence. Deuxièmement la définition d'un mécanisme d'allocation des time-slots délivrés au sein du ST en fonction des requêtes passées serait également nécessaire. Enfin du côté NCC, des algorithmes d'attribution des time-slots au sein d'une supertrame en fonction des CRs devraient être proposés et les capacités allouées devraient être contraintes par un SLA contracté entre l'INAP et le FAI pour un ST spécifique.

II.5 Conclusion

Ce chapitre a offert un état de l'art sur les normes DVB-S et DVB-RCS et l'état d'avancement de l'intégration d'IP et de la QoS dans les réseaux satellites.

Dans un premier temps, nous avons étudié la norme DVB-S et référencé les différentes méthodes d'encapsulation d'IP proposées d'une part par la norme DVB-S et d'autre part par le groupe de travail de l'IETF IP over DVB.

Dans un deuxième temps, nous avons détaillé la norme DVB-RCS dédiée à l'ensemble des mécanismes permettant à un ST d'émettre des données vers le satellite. Elle spécifie ainsi des techniques de modulation et de codage, des méthodes d'accès et des mécanismes d'allocation de bande passante.

Enfin, dans une dernière partie, nous avons réalisé un état d'avancement sur les mécanismes de QoS existants dans les systèmes satellites actuels. La conclusion que nous en avons tiré est sans équivoque. Bien que des mécanismes efficaces de gestion de ressource soient disponibles sur

la voie de retour, la plupart des réseaux satellites actuels ne les utilisent pas de manière optimale. Cela s'explique notamment par l'absence d'une architecture unifiée assurant la coordination entre les mécanismes implantés dans la couche IP et la couche MAC.

De ce fait, les systèmes DVB-RCS/S apparaissent comme les premiers systèmes satellites en mesure de fournir des services bidirectionnels multimédias large bande à un coût abordable dans les zones dépourvues d'infrastructures terrestres. Cependant, la norme DVB-RCS reste encore largement ouverte sur les questions d'optimisation de la gestion des ressources de la voie de retour et ne tire pas parti des potentialités offertes par des mécanismes tels que le DAMA au niveau MAC.

Par ailleurs, l'architecture de QoS ne saurait en aucun cas reposer uniquement sur ces seuls mécanismes de QoS de niveau 2. Ils doivent être mis en relation avec les mécanismes de QoS de niveau IP nécessaires pour supporter de bout en bout une QoS au niveau IP et tirer parti de l'ensemble des mécanismes de QoS IP existants : le contrôle d'admission, la synchronisation entre établissement de session et provisionnement dynamique de ressources dans le réseau, et le traitement différencié des différents flux applicatifs sur le réseau satellite voire le réseau utilisateur derrière le terminal satellite. Ces interactions ne sont actuellement pas évoquées dans les normes DVB-S et -RCS.

C'est pour cela que nous nous proposons dans le chapitre 3 de spécifier une architecture de QoS pour les systèmes DVB-S/RCS complète intégrant de manière cohérente les mécanismes de QoS de niveau MAC, IP et applicatifs afin d'offrir une gestion optimale des ressources sur la voie de retour et l'interopérabilité avec les mécanismes de QoS des futurs réseaux de cœur NGN. Ces étapes sont indispensables pour faciliter l'intégration des réseaux satellites d'accès dans le cadre des NGNs comme dans le cas des réseaux d'accès UMTS [98] et ADSL [99].

Chapitre 3 : Architecture de QoS pour Systèmes DVB-S/RCS

III.1 Introduction

L'objet de notre contribution est justement de spécifier cette architecture pour optimiser l'utilisation des ressources de la voie de retour. Cette architecture s'appuiera sur une gestion dynamique des ressources possibles par les potentialités d'accès multiple et d'allocation de ressources à la demande offertes par le DAMA. Elle cherchera notamment à maximiser le nombre d'utilisateurs tout en limitant le gaspillage de ressources et respectant les niveaux de QoS négociés à travers les SLAs.

Cette architecture a été proposée et développée dans le cadre du projet IST Européen SATIP6 (2001-2004) [100] et a notamment été implémentée sur une plate-forme de démonstration émulant un système satellite DVB-S/RCS.

De ce fait, le chapitre est structuré de la façon suivante.

Dans un premier temps, nous décrirons le projet SATIP6 et le scénario de démonstration sur lequel nous nous appuyerons pour justifier l'architecture de QoS.

Dans un deuxième temps, nous détaillerons le fonctionnement de la plateforme d'émulation.

Enfin nous présenterons l'architecture de QoS DiffServ adoptée et les solutions de niveau applicatif fournies afin d'offrir une classification du trafic, un contrôle d'admission et une gestion de la bande passante plus dynamiques et plus précis basés sur des informations de niveau Application et Transport.

III.2 Le projet Européen SATIP6

Les conclusions du chapitre 2 sont sans équivoque. Leur ubiquité, leur nature diffusives et leurs gestions dynamiques des ressources placent les réseaux satellites géostationnaires comme la technologie complémentaire des réseaux terrestres dans les zones géographiques reculées ou peu densément peuplées. Cependant, afin qu'ils soient en mesure de supporter la variété des services IP large bande, ils doivent intégrer des fonctionnalités avancées en terme de QoS et d'interconnexion avec les services IP traditionnels et doivent adapter, si besoin est, les protocoles traditionnels aux caractéristiques spécifiques des réseaux satellites.

III.2.1 Les objectifs

Ainsi le projet SATIP6 avait pour but initial de faciliter l'intégration des services IP dans les réseaux satellite d'accès DVB-S/RCS. Les avancées proposées n'étaient pas cantonnées à l'architecture de QoS mais s'étendaient aux domaines de la sécurité, du multicast, des protocoles de transport et de la mobilité ainsi qu'au support d'IPv6 et à son impact sur chaque domaine. Ainsi de nombreuses améliorations ont été proposées afin d'adapter les services réseaux aux avantages (propriété de diffusion) et inconvénients (délais de propagation) inhérents aux communications par satellite : optimisation des performances de TCP sur le lien satellite par l'adjonction de Performance Enhancing Proxies [101], adaptation du protocole ARP (Address Resolution Protocol) et de PIM-SM (Multicast) sur le réseau satellite, la définition d'un plan de contrôle optimisé pour le chiffrement des communications sur le segment satellite à travers le protocole FMKE (Flat Multicast Key Exchange) [102]...

Nous nous intéresserons particulièrement à l'architecture de QoS implémentée dans SATIP6 comme point de départ de l'architecture dynamique de QoS que nous étofferons au fur et à mesure du chapitre 3.

III.2.2 L'architecture SATIP6

III.2.2.1 L'architecture générique

De nombreux scénarios à plus ou moins long terme ont été envisagés dans SATIP6 comprenant satellites transparents et régénératifs et impliquant de nombreux opérateurs de services et fournisseurs d'accès se partageant les ressources d'un même système satellite. La Figure 28 offre une vision exhaustive des interactions et configurations envisagées dans SATIP6.

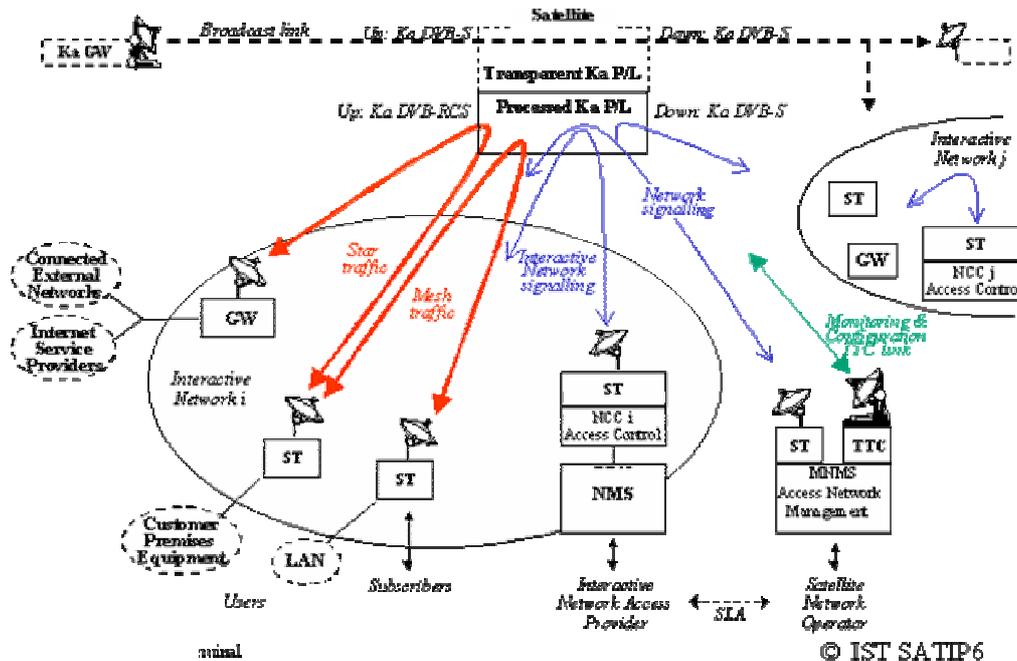


Figure 28 : Architecture générique SATIP6

III.2.2.1.1 Interactive Network

L'opérateur du réseau satellite (SNO) à travers le MNMC (Mission and Network Management Center) contrôle le système satellite en terme de position, d'alimentation, d'état de la charge utile, d'états de pertes dus aux perturbations météorologiques. Il utilise les liens TM/TC (TeleMetry/TeleCommand) pour communiquer avec le satellite. Il segmente le réseau satellite en différents Interactive Networks (INs), chacun sous la responsabilité d'un INAP. Le réseau satellite peut être partitionné en un maximum de 30 INs.

De ce fait, un IN est typiquement constitué :

- D'une station de contrôle regroupant au sein d'un ST
 - o un NCC (Network Control Center) unique centralisant la gestion dynamique des ressources
 - o un NMC (Network Management Center) s'occupant de la partie administrative
- De centaines voire de milliers de STs
- D'une Gateway (GW) fournissant l'accès vers le backbone terrestre où se trouvent les différents fournisseurs d'accès et de service.

Sans perte de généralité, nous nous cantonnerons, dans la suite du manuscrit, à la gestion des ressources dans un unique IN.

III.2.2.1.2 Capacités d'un système satellite de nouvelle génération

Afin d'avoir un ordre d'idée des ressources disponibles dans un système satellite, il faut savoir qu'un satellite actuel peut transporter quelques dizaines de transpondeurs. Astra 1H [103] est doté

de 32 transpondeurs de 24MHz/32MHz se partageant ainsi jusqu'à 1GHz de bande passante. Chaque transpondeur offre de 25 à 30Mbps ce qui fait approximativement 750 à 900 Mbps pour le satellite complet. Ces satellites sont cependant majoritairement transparents et utilisent peu de faisceaux étroits (8 pour Astra 1h).

Pour les satellites régénératifs, les capacités sont équivalentes. AmerHis [104] possède une charge hybride et met à disposition 4 canaux interconnectables de 36MHz chacun pour une capacité totale de 174Mbps. Un transpondeur DVB-RCS supporte jusqu'à 64 porteuses de 0.5Mbps chacune mais elle accepte des porteuses de 0.5, 1, 2, 4 et 8 Mbps qui peuvent être combinées au sein du même canal MF-TDMA. Les transpondeurs DVB-S offrent une capacité de 54 Mbps. Ces capacités plus modestes s'expliquent par le fait que le système AmerHis est un système « piggyback » et la fourniture de services multimédias interactifs ne constitue qu'une partie du satellite Amazonas doté lui de 51 transpondeurs.

Ensuite il nous faut considérer les capacités de transmission des STs qui sont également limitées. Selon la norme [105], les STs sont destinées à couvrir les besoins d'un simple utilisateur (144-384 kbps) à ceux d'une entreprise (2048 kbps) en passant par les accès Internet partagés en résidence (1024 kbps).

Enfin les GWs dans les configurations régénératives incorporent des STs légèrement modifiés à un coût particulièrement attractif pour les FAIs de petite et moyenne envergures et capables d'émettre en DVB-RCS au plus à 8Mbps (support de services Voix/Vidéo en plus des services d'accès à Internet).

III.2.2.1.3 Scénario d'accès et maillé

Dans la suite du manuscrit, nous différencierons le scénario d'accès et le scénario maillé. Dans le premier cas, nous considérerons un satellite transparent offrant uniquement une connectivité d'accès et, dans le deuxième cas, nous considérerons un satellite régénératif offrant à la fois connectivité d'accès et maillée. En terme de connexions, le premier scénario ne considère que des connexions à travers la GW. Ainsi pour chaque ST un seul canal MAC est nécessaire : celui à destination de la GW. Dans le scénario maillé, les ressources des liaisons montantes sont partagées entre autant de canaux MAC nécessaires pour supporter l'ensemble des connexions entre un ST émetteur et les STs destinataires.

En terme de configuration, les différences entre le scénario d'accès et le scénario maillé sont la nature de la charge utile du satellite et la localisation des fonctionnalités du NCC et NMC. En effet, dans le scénario d'accès, les fonctionnalités de NCC et NMC sont regroupées au sein de la GW tandis que dans le scénario maillé, le NMC et le NCC sont regroupés au sein d'une station de contrôle délocalisée.

III.2.2.1.4 IP dédié

Dans le cadre de SATIP6, un mécanisme d'encapsulation d'IP sur DVB a été adopté afin notamment d'offrir un mode « non connecté ».

IP dédié (IPD) est un protocole de niveau 2 subdivisé en 2 sous couches :

- La couche DLC (Data Link Control) qui implémente les fonctions liées au relais des paquets : Segmentation et Réassemblage (SAR) et résolution d'adresse
- La couche MAC (Medium Access Control) qui gère les fonctions traditionnelles d'accès (DAMA, synchronisation ...)

La couche MAC IP dédié est à même de multiplexer plusieurs instances de la couche DLC. Cette architecture permet une gestion simple du partage des ressources d'un ST pour les accès collectif.

IP dédié a été conçu pour supporter à la fois les bursts de type ATM et MPEG2-TS sur la voie retour. Sur la voie aller, c'est l'encapsulation MPE qui est utilisée.

Le format de l'entête IP dédié nous est donné par la Figure 29 :

- Label Dest : il identifie la destination du paquet

- CoS : détermine la classe de service MAC à laquelle appartient le paquet
- ID Src : identifie la source du paquet et la session client DLC à la source
- Type of Packet : détermine la nature du paquet (Signalisation ou Donnée segmentée)
- Tail Flag : mis à 1 si le paquet contient la fin d'un paquet segmenté.

Label Dest (14bits)	CoS (2bits)	Id Src (19bits)	Type of Packet (4bits)	Tail flag (1bit)
------------------------	----------------	--------------------	---------------------------	---------------------

Figure 29 : Entête d'un paquet IP-dédié

En ce qui concerne la voie ou liaison de retour, l'encapsulation la plus prometteuse est celle basée sur ATM. De ce fait, une adaptation de l'encapsulation a été considérée, les champs AAL5 et ATM vont être utilisés pour supporter les champs IP dédié :

- La concaténation du champ GFC et VPI de la cellule UNI ATM sera considérée comme le champ Label_Dest d'IP dédié
- Le champ VCI agira comme le champ ID_Src d'IP dédié
- Le champ CoS est inclus dans le champ ID_Src. On considère que pour chaque CdS correspond une session DLC identifiée par un Src_ID
- Le champ Type of Packet d'IP dédié est donné par le bit 4 du champ Payload Type ATM
- Le champ Tail Flag d'IP dédié correspond au bit 2 du champ Payload Type ATM
- La taille du paquet IP-Dédié est incluse dans le champ Length de la queue du paquet AAL5.

Le schéma d'encapsulation est donné en Figure 30.

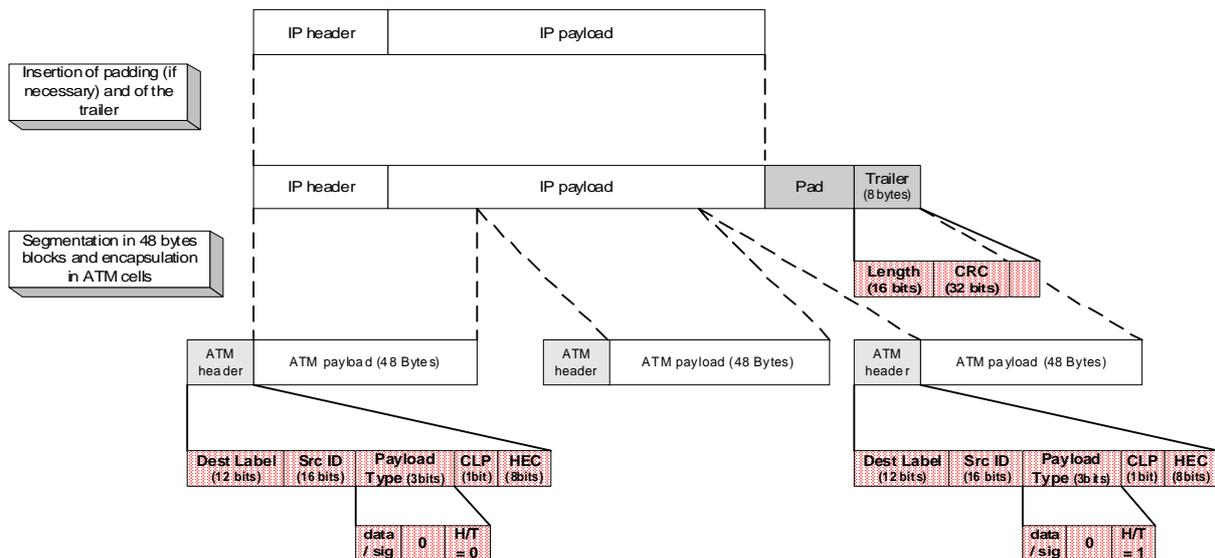


Figure 30 : Encapsulation d'IP dédié sur AAL5/ATM

III.2.2.1.5 Le protocole de résolution d'adresse satellite

Enfin le projet SATIP6 a également défini un protocole de résolution d'adresse sur satellite [107]: SARP (Satellite Address Resolution Protocol). Dans le cas exposé à la Figure 31, le protocole SARP interagit avec le protocole de résolution d'adresse IPv6 mais le protocole SARP est également défini pour fonctionner en IPv4.

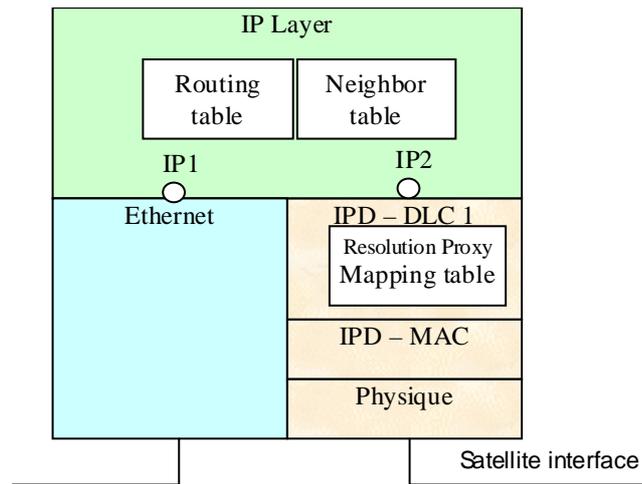


Figure 31 : Table de résolution d'adresse

Ainsi, quand la couche IPD intercepte un message de type « Neighbour Solicitation » de la couche IP, elle consulte alors sa table d'association interne (Label_Dst/adresse IP). Deux cas de figure sont alors possibles :

- si elle trouve l'adresse IP en question, elle répond par un message de type « Neighbour Advertisement » contenant le label de destination.
- sinon elle relaie le message « Neighbour Solicitation » vers un serveur central SARP localisé dans le NCC.

Ces messages sont véhiculés par l'intermédiaire du protocole C2P (Connection Control Protocol) de DVB-RCS s'appuyant sur la méthode d'encapsulation DULM. Ce protocole n'étant pas complètement standardisé par la norme, des champs sont adaptés pour pouvoir véhiculer les requêtes et réponses de type SARP.

III.3 L'architecture de QoS SATIP6

Malgré l'amélioration constante des performances de la couche physique, les ressources disponibles dans le réseau satellite sur la voie de retour notamment ne sont pas suffisantes pour permettre l'adoption d'une politique de QoS simple comme le surdimensionnement. De plus, les mécanismes de QoS dans les systèmes DVB-RCS actuels sont relativement rudimentaires et n'exploitent pas l'ensemble des possibilités offertes par le DAMA : les SLAs sont pour la plupart élémentaires, ils sont configurés statiquement et aucune signalisation dynamique de QoS n'est implémentée. Cependant l'intégration des réseaux satellites dans les NGNs ne pourra pas se faire sans l'implémentation de mécanismes de QoS avancés tels que :

- Un traitement de paquet au sein des STs et de la GW basé sur des classes de trafic
- Un contrôle d'accès aux ressources basé sur une politique de QoS
- Un contrôle d'admission et des mécanismes de réservation de ressource
- Une signalisation de QoS couplée à l'établissement de session

III.3.1 Vue d'ensemble

L'architecture DiffServ a été retenue pour ses propriétés de passage à l'échelle et pour la simplicité de son implémentation. Cette architecture entend également tirer parti des mécanismes de provisionnement dynamique de ressources et de politique de QoS récemment développés au niveau IP pour les architectures DiffServ.

Dans un premier temps, nous allons traiter dans les paragraphes suivants des mécanismes de QoS pour :

- la voie aller
- la voie retour

En effet, cette distinction est nécessaire du fait de l'asymétrie du trafic envisagé qui va entraîner une gestion de la QoS différente au niveau de la GW et des STs. Aussi, sur la voie aller, le seul et unique point d'entrée dans le domaine géré par l'INAP est la GW. Par contre, sur la voie retour, le point d'entrée est distribué entre autant de STs qui voient leur accès aux ressources contrôlé par l'algorithme DAMA.

Pour chacune des voies, nous détaillerons ses caractéristiques et l'implémentation de la QoS dans un scénario d'accès.

Enfin nous déclinerons ces mécanismes de QoS dans un scénario maillé.

Afin de simplifier la description de l'architecture, nous considérerons qu'un ST ne peut être lié qu'à un seul FAI à la fois.

III.3.2 La voie aller : la QoS dans la GW

III.3.2.1 Les fonctionnalités de la GW

Dans le scénario d'accès, la GW héberge toutes les fonctions de contrôle du NCC et de gestion du NMS. Les modules FLSS (Forward Link Sub-System) et RLSS (Return Link Sub-System) se chargent respectivement des transmissions et des fonctionnalités MAC sur la voie aller et la voie retour. Le FLSS multiplexe les paquets de données entrants et les tables SI et RCS à destination des STs dans un flux de transport MPEG2-TS. Le RLSS effectue la démodulation et le décodage des signaux DVB-RCS dont il extrait les paquets IP qu'il transmet au routeur ainsi que la signalisation DAMA destinée au NCC.

La Figure 32 détaille ainsi les fonctions supplémentaires pour le support de la QoS dans la GW. Elles comprennent notamment :

- Le routeur associé à la GW (Edge Router) qui se comporte comme un routeur de bordure interconnectant le réseau d'accès satellite et les réseaux de cœur. :
- Un module « DiffServ Server » qui rejette, lisse, marque, classe et ordonnance le trafic provenant des différents FAIs en fonction des SLAs négociés
- Le module « IP Multiplexer » (IP Mux) qui multiplexe le trafic provenant des différents FAIs à destination de la GW
- Un module « Access Gate » : pour les scénarios de QoS avancés d'établissement de session multimédia, il joue le rôle de pare-feu entre le réseau d'accès satellite et les réseaux de cœur. Il opère au niveau flux et est piloté dynamiquement par une entité externe.

Ainsi le sous-système FLSS se voit pourvu des fonctionnalités d'un routeur de cœur DiffServ à travers le module « DiffServ Server » qui applique le PHB (Per Hop Behaviour) associé à la classe de service du paquet à relayer.

Le sous-système de contrôle et de gestion regroupe les entités NMS et NCC et sont détaillés par la suite.

Le sous-système RLSS n'est quant à lui pas modifié.

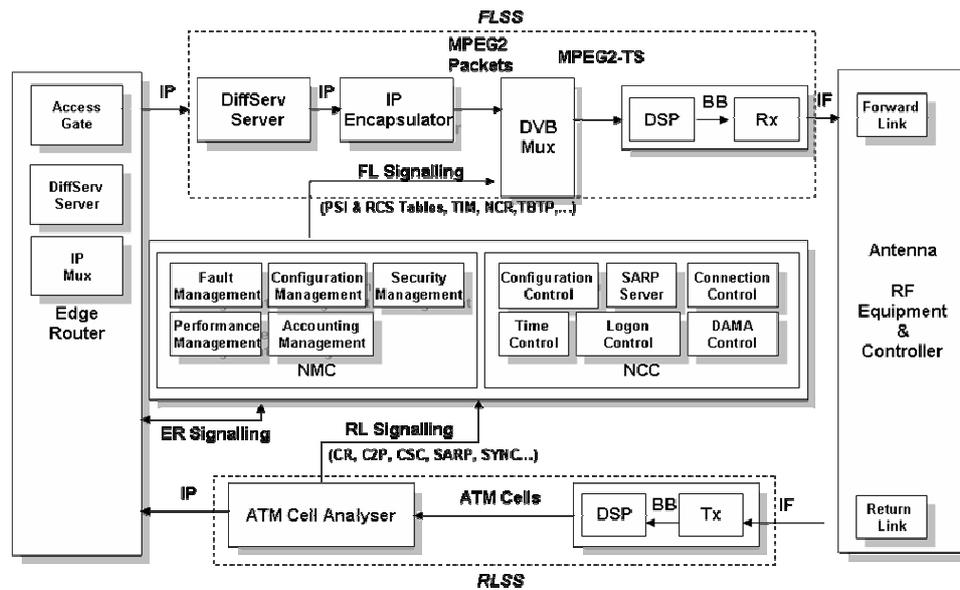


Figure 32 : Les fonctionnalités d'une Gateway

III.3.2.2 NCC et NMS

La Figure 32 donne aussi le détail des modules incorporés dans les entités NMS et NCC. Nous rappelons que le NMS fait parti du plan de Gestion et qu'il configure, supervise et collecte les données liées aux performances de la GW et des STs. Le NCC lui appartient au plan de contrôle et il alloue dynamiquement les ressources du satellite en fonction des profils de la GW et des STs (SLA notamment) contenu dans le NMS :

Le NCC est constitué ainsi des modules:

- « Configuration Control » : ce module est en charge de la configuration locale de la GW et de celle à distance des STs. Il provisionne notamment les ressources et politiques de QoS de niveau MAC et IP au sein du réseau satellite (DiffServ Server pour la GW et IP et MAC QoS pour les STs)
- « SARP Server » : ce module centralise les fonctions de résolution d'adresse (Association entre adresses IP et l'adresse MAC d'un ST)
- « Logon Control » : Ce module authentifie le ST et vérifie la validité de son compte durant la procédure de logon à partir des informations contenu dans le burst CSC. Ensuite a lieu le contrôle d'admission de la connexion d'accès associée (délégué au module « Connection Control»). L'ensemble des informations associées à ce compte stocké dans le NMS est ensuite chargé au sein du NCC. Un message TIM contenant notamment les identifiants logiques des connexions dédiées respectivement aux données d'accès et à la signalisation (VPI/VCI), le Group_id, le Logon_id, Gateway_id, Superframe_id est envoyé à destination du ST.
- « Connection Control » : en dehors de la procédure de logon, ce module est à même de modifier les paramètres des connexions existantes lorsque le ST est actif et connecté. Il peut également en ouvrir ou en fermer à la demande. Les requêtes du ST sont basées sur une signalisation de niveau MAC : le protocole C2P.
- « DAMA Control » : ce module comprend l'ordonnanceur MAC qui constitue le côté serveur des fonctions liées au DAMA. Il alloue les capacités du satellite en fonction de la politique de l'INAP et des SLAs de chaque ST et fixe notamment des capacités limites tels qu'un VBDCmin ou un RBDCmax. Ces paramètres négociés à l'abonnement du FAI sont stockés dans le NMS et peuvent être négociés dynamiquement par l'intermédiaire du protocole C2P. Ce module est à l'origine du TBTP qui donne le plan d'allocation des time-slots par

supertrame entre les STs d'un même spot montant.

- « Time Control » : Ce module gère la synchronisation des STs (procédure de synchronisation fine SYNC, grossière ACQ, et de maintenance SYNC) à travers la CMT (Correction Management Table). Il s'occupe également de la synchronisation du système satellite à travers la diffusion périodique de la NCR (Network Clock Reference).

Le NMS quant à lui gère le plan d'administration de l'IN. Il fournit toutes les informations nécessaires pour configurer et superviser les entités du réseau satellite. Il collecte également les données sur les performances de chaque équipement ce qui permet d'évaluer la fiabilité des équipements et de constituer des historiques du trafic en vue d'un meilleur dimensionnement des ressources dans le réseau.

Il provisionne les paramètres de configuration des STs, de la GW et du satellite, gère les profils des STs et le statut du compte associé en fonction de l'abonnement souscrit et du mode de facturation adopté. Il stocke ainsi l'ensemble des SLAs souscrits entre l'INAP et les FAIs ou autres fournisseurs de service que nous regroupons sous le nom de « SLA INAP-SP ». Ils les décomposent en SLAs entre opérateurs satellites (NAP) et STs regroupés sous le nom de « SLA NAP-ST ». Pour la voie aller, ces SLAs NAP-ST sont relativement simples puisqu'ils reflètent les SLAs entre FAIs et Subscribers. Par contre sur la voie retour, ces SLAs comportent des paramètres avancés liés au DAMA. La Figure 33 résume les interactions qui en découlent entre acteurs du réseau satellite.

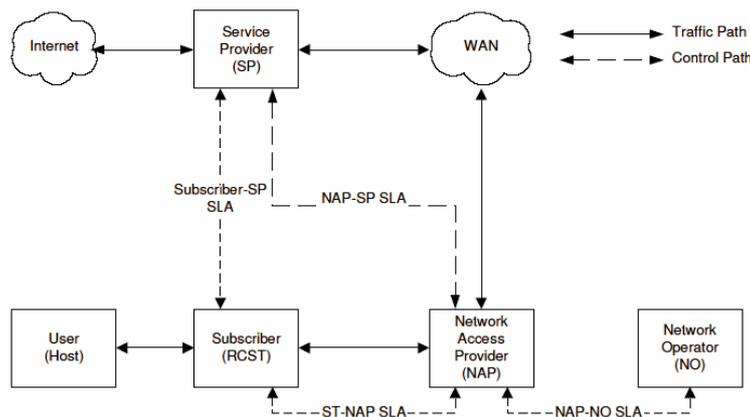


Figure 33 : Structure des SLAs

III.3.2.3 Station de contrôle en scénario maillé

Dans le scénario maillé, le satellite est régénératif et incorpore des fonctionnalités supplémentaires dont notamment la faculté de commutation à bord (commutateur, tables de commutation associées, module de contrôle de l'OBP). La Figure 34 décompose ainsi l'ensemble des fonctionnalités de niveau 2 installées dans le ST, l'OBP et la station de contrôle, ici présentée comme le NCC. Les liens entre les modules et les messages échangés sont également représentés. Nous retrouvons ainsi l'ensemble des fonctionnalités du NMC et NCC décrites auparavant en section III.3.2.2.

La différence notable avec le scénario d'accès est notamment l'apparition de la commutation (module switching) à bord du satellite. Les communications inter-STs ne nécessitent alors qu'un simple bond sur le réseau satellite. De plus, les fonctionnalités de NCC et NMC regroupées au sein de la GW dans le scénario d'accès sont regroupées dans une station à part dans le scénario maillé.

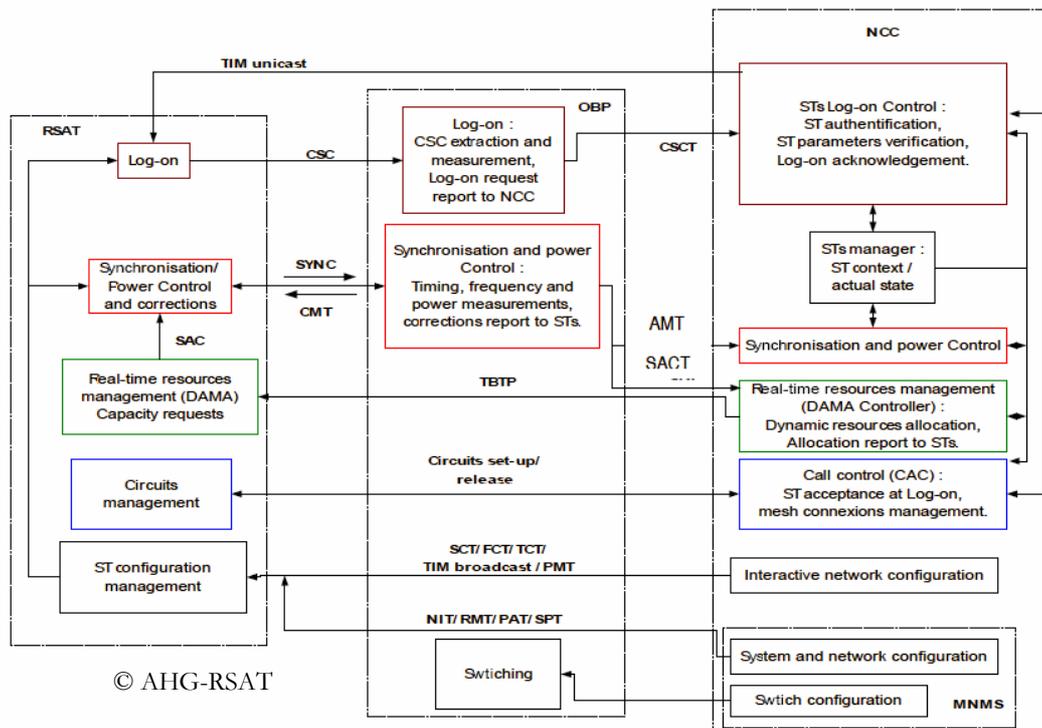


Figure 34 : Interaction entre ST, OBP et NCC dans un scénario maillé

III.3.2.4 La QoS implémentée

Les mécanismes mis en œuvre sur la voie aller sont conformes aux mécanismes traditionnels DiffServ. Les classes de service, leur niveau de priorité et leur DSCP associé sont détaillés dans la Tableau 3 et dans la référence [108].

Les classes dédiées à la signalisation sont les plus prioritaires. La signalisation est scindée en 3 catégories : Administrative, Contrôle de réseau (OSPF, DNS, DHCP, Contrôle et signalisation entre domaines administratifs) et la signalisation d'exploitation, d'administration et de maintenance (OA&M) correspondant au trafic non critique utilisant SNMP, COPS, Telnet.

Le service EF fournit, lui, un service de transfert équivalent à une ligne virtuelle dédiée : de bout en bout, délai et gigue faible, débit garanti. Les CdS AF correspondent à un même niveau de priorité mais se voient attribuées des bandes passantes différentes. Contrairement au service EF où les paquets non conformes sont rejetés, le service AF les accepte mais les marque comme des paquets non-conformes.

Trois niveaux de post-précédence sont définis et affectés au paquet en fonction de leur niveau de conformité. En cas de congestion, les routeurs intermédiaires DiffServ rejettent sélectivement les paquets selon leur degré de conformité défini.

Enfin le service BE se contente de la bande passante restante. Lors du dimensionnement des classes, il faut prendre garde à ce que la file BE ne soit pas affamée par les autres classes.

Cet ensemble de classes est exhaustif mais ne présume en rien du nombre de classes effectivement implémentées.

Tableau 3 : Classes de service et priorité inter-classe

Classes de service DiffServ	Code DSCP
Administrative, Contrôle de réseau, OAM	111000 110000 010000
EF	00101110
AF1	AF11 001010, AF12 001100, AF13 101110
AF2	AF21 010010, AF22 010100, AF23 010110
AF3	AF31 011010, AF32 011100, AF33 011110
AF4	AF41100010, AF42 100100, AF43 100110
BE	00000

La singularité de la QoS sur la voie aller qui fait d'ailleurs sa simplicité est qu'elle ne repose que sur des mécanismes de niveau IP. Comme le montre la Figure 35, les paquets provenant des ISP1 et ISP2 subissent tout d'abord un contrôle d'admission statique dans le routeur de bordure afin de vérifier que le trafic injecté par chaque ISP dans le réseau satellite est conforme avec le SLA négocié entre l'INAP et l'ISP. Le trafic de signalisation est quant à lui marqué et limité selon des règles de filtrage définies dans les SLAs INAP-SP. Les paquets sont ensuite multiplexés (IP Mux) et ordonnancés avant d'être transmis vers la GW.

Au niveau de la GW, nous retrouvons un serveur DiffServ comprenant un classifieur, un écarteur et un régulateur. Les paquets IP sont à nouveau réordonnancés car ils sont multiplexés sur la voie aller avec la signalisation DVB-S et DVB-RCS. De ce fait, un mécanisme de contrôle du débit de sortie de l'ordonnancier IP doit être mis en œuvre afin d'adapter la quantité de donnée IP acceptée au niveau MAC en fonction de la quantité de signalisation MAC générée. Nous rappelons que le débit de transmission de la GW sur la voie aller est constant. Les paquets IP sont ensuite injectés dans des canaux logiques MPEG2-TS au sein desquels les paquets resteront ordonnés. La capacité de ces canaux correspond à l'agrégation des besoins de chaque abonné en réception, par classe de service, indifféremment de l'ISP.

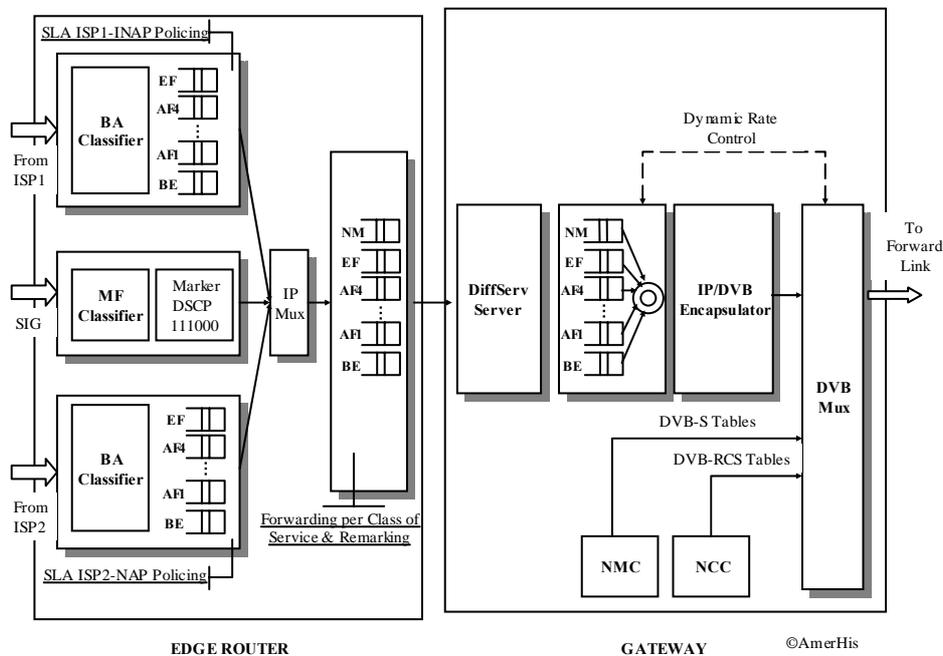


Figure 35 : QoS dans l'ER et la GW

III.3.3 La voie retour, la QoS dans le ST

Contrairement à la voie aller, les ressources d'un ST sur la voie retour sont directement partagées entre les utilisateurs. La bande passante de la voie retour est elle-même partagée entre plusieurs STs. Par ailleurs, les STs représentent des routeurs de bordure particuliers puisqu'ils sont directement connectés au segment utilisateur. De ce fait, en plus du contrôle de trafic par agrégat, un contrôle d'accès par flux fonction d'une politique de QoS et de l'état des ressources disponibles dans les classes de services DiffServ EF et AF est nécessaire. Il permet d'assurer un partage équitable des ressources entre flots et le maintien de l'intégrité du service de QoS par flot au sein d'un service DiffServ sur le segment satellite. Enfin le partage entre les STs de la bande passante liée au multiplexe MF-TDMA est assuré par des mécanismes de niveau 2.

III.3.3.1 Vue d'ensemble

La gestion des ressources satellites est effectuée à trois niveaux de granularité différents:

- Par ST au niveau MAC : les ressources de la voie retour sont partagées entre les STs. Régie par le DAMA, l'allocation des ressources est dynamique et réalise un compromis entre les besoins en bande passante et en QoS de chaque ST.
- Par agrégat au niveau IP : les ressources offertes par le FAI aux utilisateurs et s'appuyant sur les ressources fournies par la couche MAC sont réparties selon des classes de service configurées statiquement à partir de l'agrégation des SLAs souscrits par les différents utilisateurs.
- Par flux au niveau Application/Transport : L'accès aux classes de service provisionnées dans le ST doit être régie par une politique de QoS. Pour ce faire, le FAI ou le Subscriber peuvent mettre à disposition de l'utilisateur des mécanismes de signalisation et de facturation permettant dans un premier temps un accès contrôlé aux différentes classes de services selon une politique de QoS statique. Ensuite nous proposerons des solutions dynamiques permettant à l'utilisateur de modifier ces règles à la demande ² (cf. section III.5)

Et la difficulté d'une telle gestion des ressources réside dans une intégration verticale cohérente de ces trois niveaux au sein d'une seule et même architecture de QoS.

La Figure 36 détaille l'ensemble des mécanismes implémentés dans les deux premiers niveaux MAC et IP.

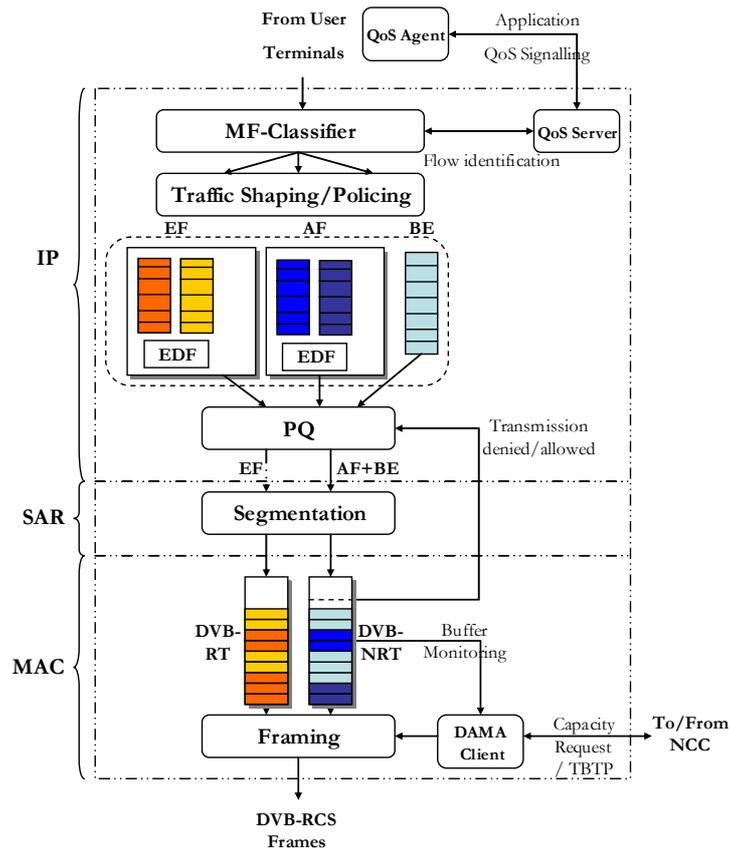


Figure 36 : Architecture de QoS d'un ST

Tout d'abord nous allons détailler le niveau MAC dans les sections III.3.3.2, III.3.3.3 et III.3.3.4. Ensuite nous passerons aux classes de service DiffServ offerte au niveau IP en section III.3.3.5. Enfin nous aborderons les interactions entre la couche IP et MAC en section III.3.3.6. Les mécanismes liés à un contrôle d'admission dynamique des flux aux classes de service IP, qui finalisent notre architecture, seront évoqués à la section III.5.

III.3.3.2 Les classes de service au niveau MAC

Le but de la QoS au niveau DVB-RCS dans le ST est d'optimiser le partage des ressources sur la voie retour entre les différents STs. Ainsi la couche MAC doit être capable à la fois :

- De fournir des garanties strictes en terme de délai et de gigue pour une minorité de flux applicatifs temps réel
- De préserver le reste des ressources satellites destiné à des applications plus tolérantes en terme de QoS en n'allouant que les ressources nécessaires

La principale difficulté au niveau MAC est donc d'offrir un algorithme de DAMA efficace. Cette efficacité passe à la fois par une judicieuse décomposition en classes de service MAC, la projection de ces classes de service sur les catégories de capacité adéquates (cf. section III.3.3.3) et enfin la pertinence des algorithmes et des informations utilisés pour requérir dynamiquement des capacités auprès du NCC (cf. section III.3.3.4).

Les deux classes de service implémentées au niveau MAC dans le terminal satellite sont :

- La classe DVB-RT (Real-Time) dédiée aux applications aux contraintes temporelles fortes (VoIP, Visioconférence)
- La classe DVB-NRT (Non Real-Time) dédiée aux applications plus tolérantes voire non affectées par le délai, la gigue et la variation de bande passante (Peer to Peer, FTP ...)

L'avantage de ces classes de service au niveau MAC est indéniablement la granularité de gestion des ressources plus fine qu'elles offrent (cellule ATM ou paquet MPEG2-TS) par rapport

au niveau IP (datagramme). Cette décomposition permet de donner la priorité à des petits datagrammes de type voix ou signalisation qui pourraient voir leur transmission retardée par de plus grands datagrammes non prioritaires. Ce gain de quelques millisecondes est essentiel dans des réseaux de type satellite. Ainsi deux classes de service par rapport aux trois classes préconisées par la norme DVB-RCS semblent suffisantes pour réaliser ce gain et permettent de ne pas trop compliquer la couche MAC.

III.3.3.3 L'association Classes de Services et Requêtes de capacités au niveau MAC

Deux modes d'allocation sont ainsi attribués à chaque classe :

- L'une que nous qualifierons de « statique » qui à la création d'une connexion assigne pour la durée de vie de la connexion une capacité fixe basée sur des capacités de type CRA
- L'autre « dynamique » régie par le DAMA développé dans le cadre du projet SATIP6 basé sur des CRs de type VBDC et sur le FCA alloué par le NCC

Afin de répondre aux exigences de la classe DVB-RT, la catégorie de capacité retenue est donc le CRA. Elle correspond à une allocation continue à chaque supertrame au logon du terminal et pour toute la durée de sa connexion. De ce fait, cette capacité est toujours disponible et n'est pas sujette à la latence introduite par le DAMA. L'inconvénient d'une telle allocation est que la capacité est fixe et peut aboutir à une sous-exploitation des ressources du satellite.

Quant à la classe DVB-NRT, elle bénéficie de capacités allouées à la fois par des requêtes de type VBDC et FCA. Les requêtes VBDC correspondent à un volume de cellules demandées au NCC calculé sur le taux d'occupation et de remplissage de la queue MAC DVB-NRT et qui sont actualisées et émises à chaque supertrame. Ce type de réservation est vital afin de préserver les ressources du satellite puisqu'elle permet d'allouer les ressources nécessaires au ST à la cellule ATM près. Cependant, entre l'émission de la requête VBDC et la consommation effective de ces capacités par le ST, le cycle de Requête/Allocation (défini dans le paragraphe suivant) ne peut descendre en deçà d'une valeur minimale appelée T_{MSL} pour Minimum Scheduling Latency (également défini dans le paragraphe suivant) et peut aller jusqu'à plusieurs secondes si une supertrame ne suffit pas à écouler le volume requis par l'ensemble des ST à un instant précis (congestion au niveau de la liaison montante gérée par le NCC). Ces requêtes étant cumulatives, les volumes accumulés au niveau du NCC sont réparties sur plusieurs supertrames jusqu'à ce que la congestion se résorbe.

Le cycle de Requête/Allocation (cf. Figure 38) est défini comme l'intervalle de temps entre le début du calcul d'une CR par un ST et le moment où ces capacités requises peuvent être utilisées par le ST. Ce cycle a un délai minimum incompressible T_{MSL} que l'on nomme MSL (Minimum Scheduling Latency). Il correspond à la somme de T_{RC-ST} temps de calcul d'une CR et de sa transmission, T_{AIR} temps de propagation sur l'interface air (du ST/NCC au satellite et inversement), T_{SAT} temps de traitement au sein du satellite (réception de la CR, constitution de la table SACT et émission de la SACT), T_{NCC} temps de traitement au sein du NCC (réception du SACT+ calcul du TBTP + émission du TBTP) et enfin $T_{TBTP-ST}$ temps nécessaire à la réception et au traitement du TBTP.

$$T_{MSL} = T_{RC-ST} + 4 \times T_{AIR} + T_{NCC} + T_{SAT} + T_{TBTP-ST}$$

Les temps généralement considérés sont de 20ms pour $T_{RC-ST} + T_{TBTP-ST}$, pour 20ms pour T_{SAT} , 125 ms pour T_{AIR} , 60 ms pour le T_{NCC} soit un T_{MSL} de 600ms.

III.3.3.4 Le DAMA

Le DAMA est un protocole de type Client/Serveur permettant dans un système DVB-RCS d'allouer à la demande des ressources de niveau MAC. Le client et le serveur sont respectivement localisés dans chaque ST et dans le NCC. La signalisation des STs vers le NCC retenue dans SATIP6 est une signalisation hors-bande. Elle est réalisée par l'intermédiaire du préfixe SAC dans les bursts SYNC. Le NCC quant à lui répond par la diffusion de la table TBTP à chaque

supertrame.

Le DAMA peut se décomposer en 6 phases distinctes :

- Le calcul des CRs dans les terminaux
- La transmission de ces CRs vers l'ordonnanceur MAC
- Le calcul par l'ordonnanceur MAC des ressources à allouer
- La répartition de ces ressources au sein de la supertrame
- La génération et la transmission du TBTP définissant le plan d'allocation de time-slots
- La distribution à l'intérieur des STs de ces ressources entre les utilisateurs finals et leurs applications

Actuellement le caractère particulièrement ouvert de la norme DVB-RCS sur le DAMA favorise les solutions propriétaires difficilement accessibles. De fait, la plupart des évaluations actuelles de performance sur satellite sont basés sur des DAMAs certes réalistes mais fortement liées au modèle de trafic adopté. Ainsi, par exemple, dans [111] et [112], le délai d'accès induit par le protocole de contrôle d'accès est approximé par un délai fixe donc indépendant des caractéristiques statistiques du trafic et des conditions de charge du réseau. Des protocoles plus avancés tels que [113] proposent un algorithme se basant sur la prédiction de bursts de trafic. Cependant il se base sur des profils de trafic de type On-Off alternant des périodes d'activité et de silence qui ne sont pas représentatifs de l'ensemble des services multimédias.

L'algorithme proposé dans SATIP6 s'appuie sur des outils de la théorie de la commande. L'allocation des ressources est basée non pas sur un modèle de trafic particulier mais sur des mesures de trafic en temps réel.

III.3.3.4.1 La structure de la supertrame

Afin d'introduire la formule de l'algorithme utilisée dans les STs de SATIP6 pour calculer les requêtes VBDC, nous allons préciser le format de la supertrame illustré par la Figure 37.

Cette supertrame est décomposée en plusieurs trames et les trames de signalisation (SYNC, CSC, ACQ) sont regroupées en début de supertrame. Le mode de partage MF-TDMA adopté est statique ce qui fixe la taille et le débit des time-slots utilisés par les STs. On considère également que les bursts ATM ne contiennent qu'une unique cellule ATM.

Dans notre plateforme de démonstration, n (le nombre de trames de donnée constituant une supertrame) est fixé à 10. Ainsi chaque supertrame est constituée d'une première trame de signalisation dédiée notamment aux CRs et de 10 trames dédiées au trafic.

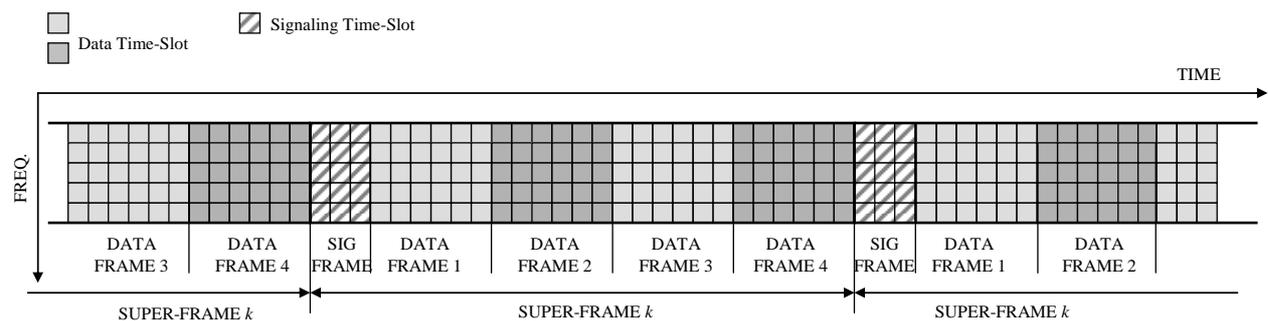


Figure 37 : Structure d'une supertrame composée de $n=4$ trames

La durée de la supertrame est de 500ms qui constitue également la période à la fois de rafraîchissement du TBTP et des CRs pour un ST.

III.3.3.4.2 L'algorithme de calcul des requêtes de capacités

Le DAMA doit réussir le compromis entre deux objectifs divergents :

- Une utilisation optimale des ressources de la voie montante : les STs doivent être en mesure de consommer la totalité des capacités qu'ils ont requise. Cela signifie qu'une quantité appropriée de paquets doit être accumulée au niveau des files MAC. En effet si les STs ne disposent pas de cellules ATM dans leurs buffers durant une supertrame, l'ensemble des time-slots qui leur étaient dédiés pour cette supertrame est perdu.
- La latence due au DAMA : la taille de la queue MAC doit être réduite afin de diminuer le temps de séjour des cellules dans la queue MAC. Ainsi à la durée du cycle Requête/Allocation, il faut adjoindre le temps d'attente dans la queue DVB-NRT avant une opportunité de requête et le temps d'attente avant la transmission effective de la cellule ATM.

Ces deux critères antagonistes sont les principaux indicateurs de performance d'un protocole DAMA

Le cycle Requête/Allocation est illustré par la Figure 38. Il considère des trames de 50ms et une supertrame composée de 10 trames. Les slots de signalisation n'apparaissent pas pour ne pas surcharger la figure. Les CRs sont émis dans les slots de signalisation en début de supertrame. Les TBTPs sont également envoyés toutes les 500ms. Ici on suppose que le ST se voit attribuer l'ensemble des ressources qu'il a exigé dans le prochain TBTP (Pas de congestion au niveau du lien montant géré par le NCC).

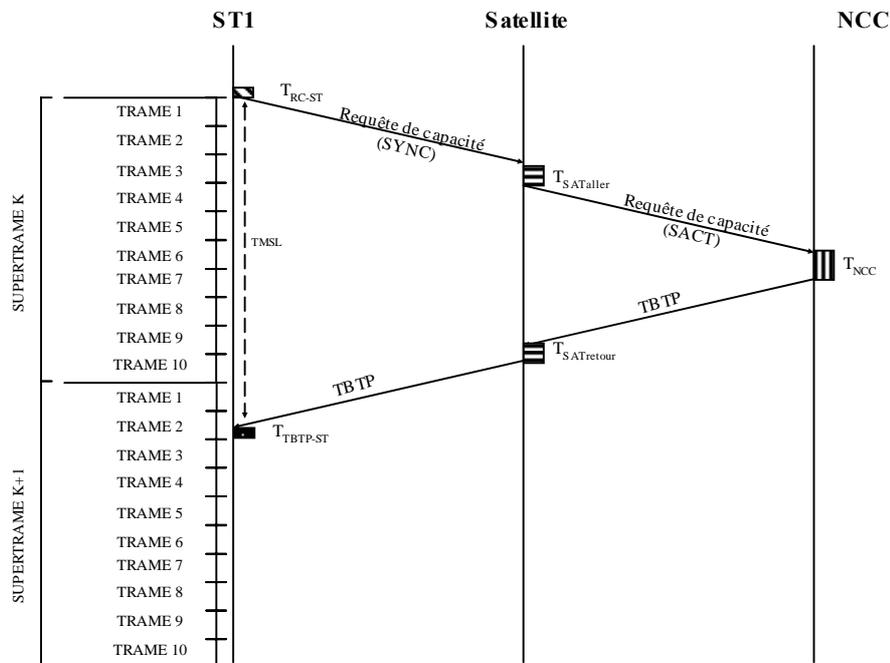


Figure 38 : Cycle de Requête/Allocation DAMA

A la $k^{\text{ième}}$ supertrame, le ST calcule la CR qu'il effectuera selon la formule suivante [109]:

$$r_{REQ}[k] = \left[r_{IN}[k] + \max \left\{ 0, \frac{1}{2} [q[k] - (r_{IN}[k] + (\alpha - 1)r_{IN}[k-1] + (\alpha - 1)r_{IN}[k-2] + r_{REQ}[k-1] + r_{REQ}[k-2])] \right\} \right]$$

où $r_{REQ}[k]$ est la capacité requise à la $k^{\text{ième}}$ supertrame, $r_{IN}[k]$ est la capacité correspondant aux nombres de cellules reçues durant la $(k-1)^{\text{ième}}$ supertrame, $q[k]$ est la taille de la queue MAC DVB-NRT mesurée au début de la $k^{\text{ième}}$ supertrame et α le coefficient d'anticipation.

Le premier terme de la formule correspond aux capacités requises pour écouler les cellules ATM qui sont entrées dans la queue DVB-NRT durant la supertrame précédente. Afin de se prémunir contre une famine éventuelle due à une congestion au niveau du lien montant géré par le NCC, le second terme est proportionnel à la taille de la queue estimée à la supertrame $k + T_{MSL}$.

(moment où l'allocation de ressource correspondante sera reçue) ; cette estimation est calculée à partir de la taille actuelle de la queue moins la somme des requêtes encore en attente de réponse. De plus, le deuxième terme est également proportionnel à une somme de cellules entrantes sur les supertrames passées. La part de ce dernier facteur est néanmoins pondérée par un facteur α , une constante entre 0 et 1. Ainsi en ajustant ce coefficient α , le temps de séjour dans la queue MAC peut être réduit. Cette réduction entraîne en contrepartie une surallocation de ressources ce qui tend à diminuer l'efficacité d'utilisation du médium.

La Figure 39 résume les résultats d'une campagne de simulation [111] menée afin d'évaluer notamment l'influence du paramètre α sur la latence et l'efficacité de l'algorithme. Comme l'on pouvait s'y attendre, l'efficacité croît avec α tandis que le délai moyen d'attente diminue avec α .

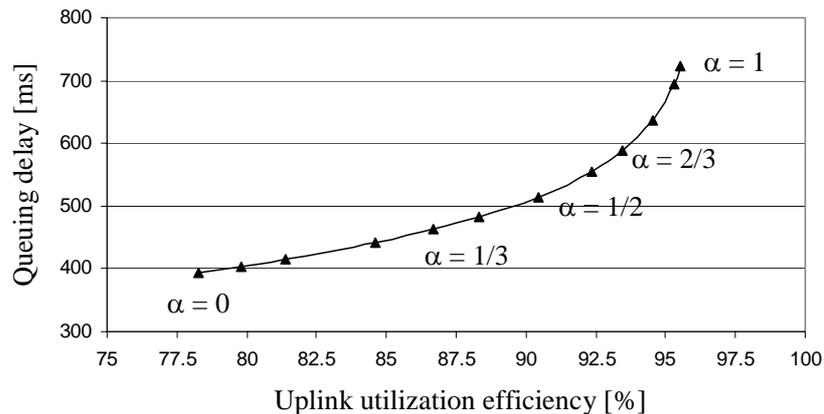


Figure 39 : influence d' α sur l'efficacité du DAMA

Pour plus de détails, le lecteur pourra se référer aux articles suivants [109][110].

Comme nous pouvons le voir sur la Figure 36, un mécanisme d'interaction entre la couche MAC et la couche IP a été également introduit afin d'éviter notamment les phénomènes de congestion au niveau MAC. Ce mécanisme consiste en un seuil au niveau de la queue MAC DVB-NRT qui, s'il est dépassé, empêche l'ordonnanceur de niveau IP d'envoyer des paquets supplémentaires vers la file DVB-NRT tant que sa taille n'est pas repassée sous ce seuil. La question est alors de savoir quelle est la valeur optimale de ce seuil. Il est directement calqué sur la taille maximale de la queue obtenue en l'absence de congestion du lien montant. Ce cas de figure est illustré dans la Figure 40.

Dans la configuration de notre plateforme d'expérimentation, nous avons retenu une supertrame de 500 ms décomposée en 10 trames de 50 ms. La période des CRs est confondue avec celle des supertrames et T_{MSL} est supérieur à la durée d'une supertrame ; donc, au pire, les cellules peuvent s'emmagasiner pendant 3 périodes avant de recevoir un TBTP lui permettant d'écouler les premières cellules emmagasinées. Ce pire cas est illustré dans la Figure 40 où la taille de la queue MAC DVB-NRT est représentée en ordonnée en fonction du temps. Découpé en 2 phases, on a tout d'abord la phase 1 où les cellules s'accumulent à un débit constant, le débit maximum du ST. Ensuite on passe en régime stationnaire où le débit d'entrée est égal au débit sortant. On remarque que pendant 3 supertrames aucune cellule n'est envoyée : la taille maximum de la queue MAC DVB-NRT est alors atteinte.

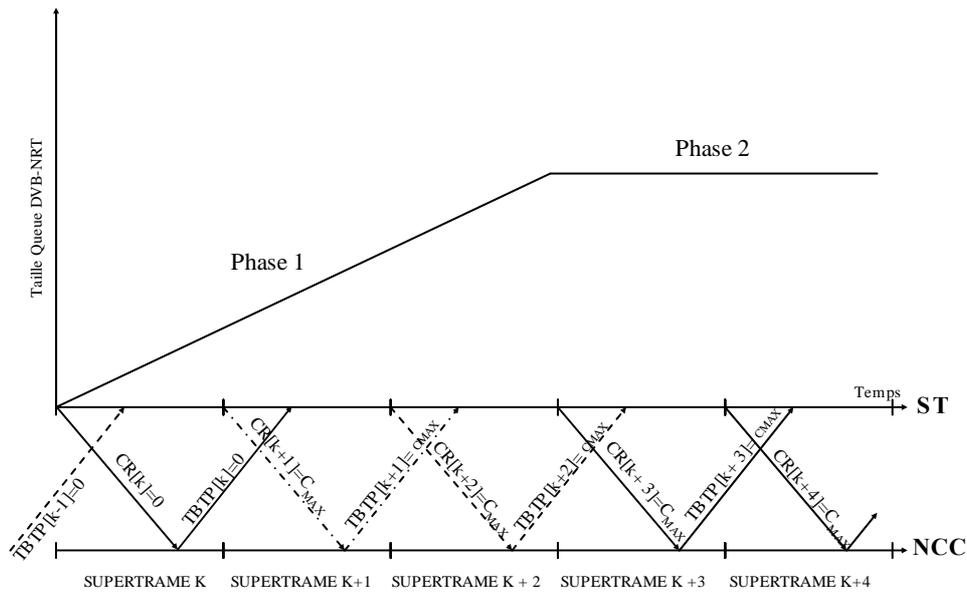


Figure 40 : Saturation de la queue MAC DVB-NRT

Le seuil est donc fixé par la formule suivante :

$$Threshold = r_{MAX} * (\Delta t) \text{ avec } r_{MAX} \text{ le débit du ST et } \Delta t \text{ les 3 périodes considérées.}$$

Nous noterons que ce calcul du seuil n'est valable que dans le cas de figure où la période des CR et des supertrames est supérieure à T_{MSL} .

Bien qu'élevé, la diminution de ce seuil pourrait entraîner des pertes de cellules ATM par débordement ce qui aurait pour résultat une transmission partielle d'un PDU AAL5 et donc un gaspillage de ressource. S'il était augmenté, le séjour des cellules ATM dans la couche MAC serait inutilement allongé.

III.3.3.4.3 L'ordonnanceur MAC

L'ordonnanceur MAC constitue la partie serveur du DAMA et il est en charge du traitement des CR et de l'allocation des ressources correspondantes par supertrame à travers le calcul du TBTP. Il est à même de différencier les trois catégories de capacités référencées précédemment qui sont traitées selon leur niveau de priorité : $CRA > VBDC > FCA$.

Ainsi l'ordonnanceur MAC alloue tout d'abord les capacités correspondantes au CRA de chaque ST connecté. Ensuite, il traite les requêtes de VBDC selon un algorithme de partage équitable. Du fait qu'elles sont cumulatives, le compte de chaque ST est crédité à la réception de chaque nouvelle requête VBDC. Pour le calcul du TBTP, chaque compte est débité tour à tour jusqu'à ce que l'ensemble des comptes affiche un solde nul ou que les capacités totales de la supertrame aient été allouées (congestion).

Enfin si toutes les capacités n'ont pas été allouées, les time-slots restants sont distribués selon un algorithme de Round Robin à hauteur d'un seuil maximum fixé dans le NCC. Cette dernière catégorie d'allocation est le FCA, à la charge du NCC, qui permet aux STs, dans des conditions de charge très faible, de bénéficier d'un minimum de cellules allouées sans avoir à attendre au minimum T_{MSL} pour être capable de les transmettre.

III.3.3.5 QoS niveau IP

Dans un souci de résistance au facteur d'échelle, l'architecture de QoS proposée au niveau IP sépare le trafic en trois classes de service :

- BE ne garantissant rien
- AF assurant une QoS relative
- EF garantissant une QoS absolue.

Le classificateur est de type multi-champs : il se base sur une liste de quintuplets <IP Src, IP Dst, Port Src, Port Dst, type de protocole> permettant d'identifier un flux applicatif pour l'orienter. Ce type a été retenu car la majeure partie des applications courantes n'est pas en mesure de marquer le champ DCSP de leur paquet.

Le cœur de l'architecture de QoS IP est basé sur une politique d'ordonnancement basée sur la discipline EDF (Earliest Deadline First) pour l'ordonnancement et des Token Buckets (RC-EDF, Rate Controlled-EDF [116]) pour le contrôle du trafic AF et EF. Les principaux avantages de cette approche sont le coût minimal de calcul (l'algorithme EDF a un coût logarithmique) et un contrôle strict du délai et de la bande passante. Ainsi les Token Buckets garantissent une allocation minimale de bande passante, l'isolation des flux et une équitable répartition des ressources entre les flux. Les trafics EF et AF sont gérés par des ordonnanceurs EDF séparés tandis que le trafic BE est lui stocké dans une simple file FIFO. Les trois catégories sont ensuite servies par un ordonnanceur de type PQ (Priority Queuing). Sachant que le trafic EF est « mappé » sur la CdS DVB-RT et que les trafic AF et BE se partagent la CdS DVB-NRT, les paquets IP sont extraits vers la couche MAC de la manière suivante :

- Les paquets EF réordonnés à travers une discipline EDF sont servis tant que les files EF sont pleines.
- Les paquets AF sont ensuite servis et orientés vers la queue MAC DVB-NRT
- Le trafic BE est enfin passé directement vers la queue MAC DVB-NRT quand les files AF et EF sont complètement vides

En ce qui concerne l'affiliation entre les classes de service et les applications, nous invitons le lecteur à se référer aux classes de trafic BSM déjà évoquées à la section I.3.3.1.1.3.

Un contrôle d'admission statique au niveau des services AF et EF est là pour assurer que la file BE ne soit pas affamée et qu'un minimum de bande passante lui soit alloué.

III.3.3.6 Interaction IP/MAC

Un certain nombre d'interaction entre les couches IP et MAC sont nécessaires afin d'optimiser le fonctionnement de cette architecture. Le seuil MAC pour la file DVB-NRT évoqué en section III.3.3.4 met en œuvre ainsi un mécanisme de contrôle de la couche IP pilotée par des informations de la couche MAC.

D'autres précautions sont également à prendre en ce qui concerne l'ordonnancement et le service des cellules ATM et des paquets IP en terme de préemption.

Tout d'abord, il faut s'assurer qu'aucun entrelacement de cellules ATM provenant de paquets IP différents ne soit possible au sein d'un PVC. Ainsi l'arrivée d'un paquet AF ne pourra en aucun cas préempter la transmission d'un paquet BE.

Par ailleurs, le trafic EF bénéficie de son propre PVC donc le trafic EF peut préempter le trafic AF/BE au niveau ATM. Ainsi, au niveau MAC, la transmission d'une cellule ATM d'un paquet BE ou AF partiellement transmis pourra être préemptée par un paquet EF. A la prochaine trame, les cellules issues du paquet EF sont prioritaires.

En ce qui concerne la calibration des paramètres du module « Trafic Shaping/Policing », elle doit être mise en relation avec un certain nombre de paramètres de la couche MAC, notamment le CRA, $VBDC_{min}$ et $VBDC_{max}$ et le débit de transmission maximum du ST comme nous le verrons dans la présentation de la plateforme d'expérimentation.

III.3.3.7 Le scénario maillé

Quelles sont les modifications à apporter à notre architecture de QoS dans un scénario de type maillé?

Le satellite est doté d'une charge utile régénérative afin de pouvoir commuter les cellules selon les circuits mis à jour dans sa table de commutation par le NCC. Le NMC et NCC sont découplés de la GW et sont réunis au sein d'une station indépendante (cf. Figure 41).

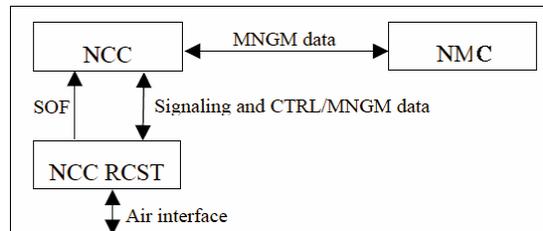


Figure 41 : Station de contrôle regroupant NCC et NMC

Les seules fonctions de contrôle à la charge de l'OBP sont d'extraire et de concaténer les CRs dans des bursts SYNC dans une table SACT envoyée vers le NCC.

Dans ce nouveau scénario, nous considérerons que les STs disposent toujours par défaut d'une connexion d'accès à travers une GW-ST unique dans l'IN, connexion qui est établie au logon. Un canal de signalisation est également ouvert vers la station de contrôle si elle est localisée dans un spot différent de celui du ST. Enfin des connexions supplémentaires vers d'autres spots peuvent être déclenchées au logon du ST pour la mise en place de VPNs par exemple ou dynamiquement par le protocole C2P vers les spots où se trouvent les STs avec lesquels on souhaite communiquer directement (cf. Figure 34). Le contrôle d'admission des connexions dans le NCC déterminera si la connexion requise peut être admise dans un canal existant avec ou sans reconfiguration de ce dernier ou bien si elle nécessite la création d'un nouveau canal.

Dans le scénario d'accès, le trafic provenant d'un ST n'a qu'une seule destination possible : la GW. Dans le scénario maillé, le ST peut disposer de plusieurs canaux, chacun associé à un spot de destination distinct. Afin de gérer de manière indépendante le trafic de chaque canal, l'architecture de QoS présentée dans le scénario d'accès doit être répliquée pour chaque canal. Ainsi le ST émet des CRs différentes pour chaque canal et l'ordonnanceur MAC gère les CRs de chaque canal indépendamment.

III.3.3.8 Conclusions

En comparaison avec les architectures traditionnelles qui reconditionnent et réordonnancent les cellules ATM au niveau MAC, l'architecture SATIP6 présente de nombreux avantages :

- Elle évite la duplication des mécanismes d'ordonnancement et de conditionnement au niveau MAC et IP
- Elle évite le débordement des files MAC et par conséquent la transmission partielle d'un paquet IP qui entraîne un gaspillage de ressource
- En cas de congestion, les paquets accumulés le sont dans les files IP et non les files MAC et bénéficient par la même des capacités de discrimination de la politique d'ordonnancement qui privilégie les paquets avec les contraintes en terme de QoS les plus fortes

Bien que les mesures soient effectuées à partir d'informations de niveau MAC pour des raisons de simplicité d'implémentation, nous remarquerons que l'utilisation de mesures au niveau des files IP permettrait de réduire le seuil MAC à un niveau minimum. La queue MAC ne serait alors utilisée que pour emmagasiner les paquets destinés à la nouvelle supertrame rendant ainsi le

temps de séjour dans les files MAC négligeable.

L'architecture SATIP6 remplit ainsi les objectifs qu'elle s'était fixée. Ainsi cette architecture dote les réseaux satellites d'une architecture DiffServ qui leur permet de supporter un traitement différencié des paquets au niveau IP. De plus, l'architecture proposée couple ces mécanismes avec des méthodes d'accès aux ressources satellites de niveau 2 basées, soit sur un contrôle d'admission orienté connexion comme le CAC (Connection Admission Control), soit sur une allocation dynamique à la demande (DAMA). Ces mécanismes sous spécifiés dans la norme DVB-RCS sont ici pleinement détaillés.

Afin de pouvoir évaluer l'efficacité de ces différents mécanismes, cette architecture a été implémentée dans une plateforme d'émulation que nous allons décrire dans la section suivante.

III.4 La plateforme expérimentale SATIP6

Le projet SATIP6 avait pour objectif la constitution d'une plateforme d'émulation d'un système satellite DVB-S/RCS implémentant l'architecture de QoS détaillée précédemment afin d'être en mesure de l'évaluer de bout en bout à partir d'applications réelles.

III.4.1 Le scénario de démonstration

Le scénario retenu pour les besoins de la démonstration est illustré dans la Figure 42. Il consiste en l'interconnexion d'une société mère et de ses deux succursales à travers un satellite géostationnaire régénératif DVB-S/RCS. Ce système s'appuie sur des liaisons montantes Ka MF-TDMA et descendantes Ku TDM. La GW et les STs sont les routeurs d'entrée/sortie du domaine satellite équivalent à un AS DiffServ.

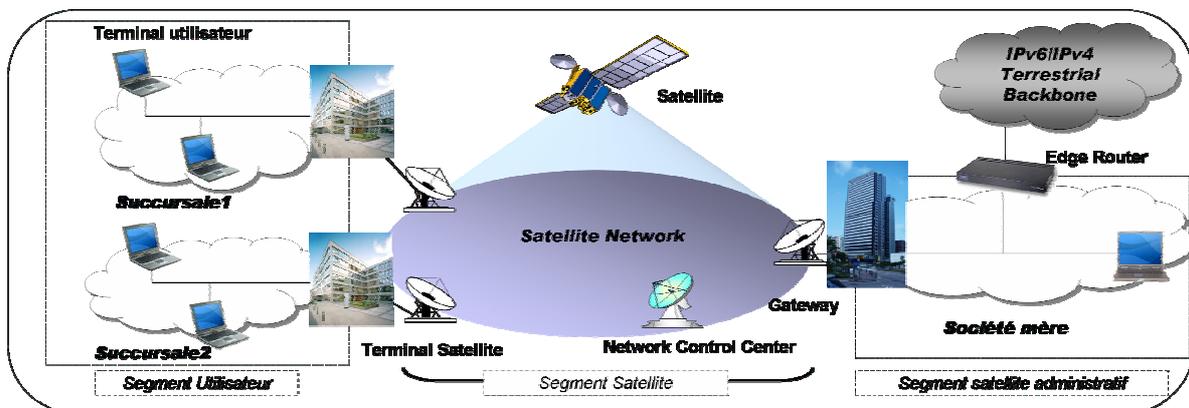


Figure 42 : Scénario de démonstration SATIP6

A gauche de la Figure 42, nous retrouvons le segment utilisateur de la plateforme. A droite, nous avons le réseau de la société mère comprenant la GW et l'accès aux réseaux terrestres. Au centre, se trouve le segment satellite émulé.

Le scénario de base retenu pour l'évaluation de la plateforme dans le chapitre 4 est le suivant :

- Satellite régénératif monospot : les spots montant, descendant et de signalisation sont confondus
- Les entités satellites sont configurées manuellement par l'intermédiaire de fichiers de configuration
- La table de commutation embarquée et le cache de résolution d'adresse des STs sont configurés statiquement
- La station de contrôle se résume au NCC qui se contente de gérer les procédures de logon, la synchronisation temporelle, et le DAMA. Tout ce qui est relatif à l'authentification, l'autorisation et les aspects de facturation, pertinents dans des conditions d'exploitation, n'est

pas implémenté dans la plateforme expérimentale.

Toutefois la plateforme est en mesure d'émuler une couverture multispot et propose également un mode d'émulation d'un satellite transparent.

Pour ce scénario de base, la plateforme est composée de trois réseaux Ethernet interconnectés par un réseau satellite émulé. Dans cette configuration, un minimum de sept stations est nécessaire pour réaliser les scénarios « unicast ». Le réseau satellite comme nous allons le voir est émulé sur un réseau Ethernet 100 Mbps. La raison de ce choix ainsi que les fonctionnalités implémentées dans les différentes entités du réseau satellite sont détaillées dans les paragraphes suivants.

III.4.2 La plateforme d'émulation

Afin de bénéficier d'une plateforme évolutive et modulaire, des contraintes fortes ont été fixées lors de son développement.

III.4.2.1 Le système d'exploitation support : Linux

L'un des premiers impératifs était de profiter au maximum des fonctionnalités du système Linux qui supporte nativement IPv6 depuis la distribution Red Hat 9.0 mais aussi un ensemble d'applications IPv6 (Apache comme serveur HTTP, Mozilla comme client HTTP, Vsftpd comme serveur FTP, Gnomemeeting pour la vidéoconférence, VideoLanClient pour le streaming de vidéo).

III.4.2.2 L'environnement d'émulation

Le deuxième impératif est la modularité de l'émulation. Afin d'offrir une structure modulaire, la partie émulation satellite a été réalisée dans un environnement spécifique de développement. Le runtime utilisé s'appelle Margouilla [118] et est basé sur une license LGPL (Library General Public License). Il a été développé en C++ et fournit un environnement de développement et d'exécution indépendant du système d'exploitation en définissant un gestionnaire d'évènement complet comprenant échange de messages, threads, timers, évènements Socket et traces. Il permet ainsi de spécifier complètement un protocole et d'en évaluer les performances. De nombreux protocoles classiques sont déjà implémentés (IP/ATM/Ethernet ...). De plus, un éditeur graphique permet de décrire les blocs composant un système et leur interaction selon le standard graphique SDL (Specification and Description Language). Du code source peut être généré directement à partir de ces modèles facilitant le travail du développeur. Enfin, des fonctionnalités avancées permettent la génération de traces et de messages d'erreur complets lors de l'exécution d'un protocole spécifié.

Les blocs Margouilla développés pour les besoins de la plateforme SATIP6 font partie de la pile protocolaire détaillée dans la Figure 43. Il s'agit des blocs :

- *satellite carrier* responsable de l'émulation des différentes porteuses sur Ethernet et de l'émulation du délai et des erreurs du lien satellite
- *DVB-S/RCS* qui formate les données selon la norme DVB-S/RCS et remplit les trames DVB-RCS avec les cellules ATM provenant de la couche IP-dédié sous le contrôle du bloc *DAMA*
- *DAMA* qui implémente les algorithmes DAMA permettant de gérer les ressources satellites à la couche 2
- *IP Dedicated* responsable des fonctions de segmentation et de réassemblage. Il implémente également le mécanisme de résolution d'adresse
- *IP QoS* qui implémente les fonctions de différenciation et de conditionnement de niveau IP

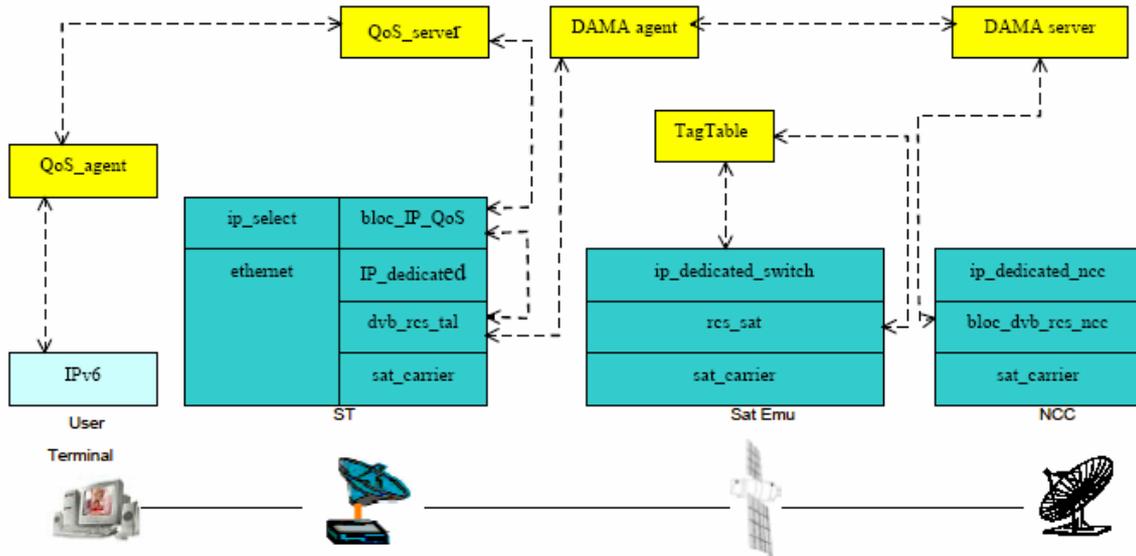


Figure 43 : Pile protocolaire SATIP6

III.4.3 L'émulation de la couche physique du segment satellite

III.4.3.1 L'émulation des liaisons montantes et descendantes

L'émulation des liaisons montantes et descendantes a été réalisée au dessus d'Ethernet. Ethernet a été choisi pour ses propriétés diffusives et pour ses capacités en terme de bande passante. Ainsi chaque trame DVB est encapsulée dans une trame Ethernet brute. Chaque canal satellite correspondant à une porteuse est associé à une adresse multicast.

Pour chaque spot, on distingue ainsi quatre canaux :

- Un canal dédié aux données DVB-S sur une liaison descendante
- Un canal dédié aux données DVB-RCS pour une liaison montante
- Deux canaux dédiés à la signalisation de contrôle et de gestion (Connection request, TBTP, Connection Confirm, SACT...), un pour la liaison montante et un pour la liaison descendante

L'émulateur satellite écoute l'ensemble des adresses multicast correspondant aux canaux ascendants et chaque ST écoute l'adresse multicast correspondant au canal descendant de son spot. Comme nous l'avons déjà dit précédemment, une adresse multicast est associée à chaque canal. Une discrimination de plus haut niveau est nécessaire pour filtrer les informations qui ne sont pas destinées à un ST. Ainsi la table SACT reçue par les STs non NCC doit être rejetée.

Ainsi, sur le réseau Ethernet 100 Mbps, le trafic de données unicast de chaque ST est dédoublé sur le réseau Ethernet puisque le trafic entrant dans le satellite à destination de l'adresse multicast de la liaison montante est réémis avec l'adresse de destination multicast de la liaison descendante. Pour le trafic multicast, la capacité du trafic montant est multiplié par autant de spots descendants concernés.

Ainsi avec trois STs situés dans trois spots descendants différents et émettant chacun des flux multicast de 8Mbps à destination des deux autres spots, le réseau Ethernet serait alors chargé à 72 Mbps. Ces débits d'émission sont largement supérieurs à ceux que nous allons envisager dans nos scénarios d'évaluation. On peut donc considérer qu'un réseau Ethernet de 100 Mbps est suffisant.

Par ailleurs, une adresse multicast Ethernet est entièrement dédiée aux messages de synchronisation. Des trames spéciales sont ainsi émises directement par le NCC vers les STs toutes les supertrames ce qui permet aux STs de se resynchroniser régulièrement. Entre deux

supertrames, un mécanisme de synchronisation interne est lancé dans chaque ST qui provoque un réveil périodique (toutes les trames) du processus émettant les trames DVB-RCS.

III.4.3.2 L'émulation de lien satellite

L'émulateur satellite a été développé dans le cadre du projet IST Brahms [117] et peut prendre de multiples facteurs en compte. Ainsi le SE (Satellite Emulator) simule différents systèmes satellites en permettant la configuration du délai, de la gigue, du profil des pertes, du type de satellite et du type de connectivité.

Les systèmes satellites émules vont de simples satellites transparents GEO ou non GEO monospot ou multispots à des satellites GEO régénératifs avec commutation de circuit par label.

Le SE simule un lien satellite bidirectionnel en temps réel. Les paquets envoyés par un ST sont emmagasinés dans des buffers au sein du SE, ensuite commutés puis subissent un retard et des séquences d'erreur bit conformément à la configuration du SE avant d'être envoyés vers la liaison descendante émulée. La Figure 44 nous indique le trajet d'un paquet dans le SE.

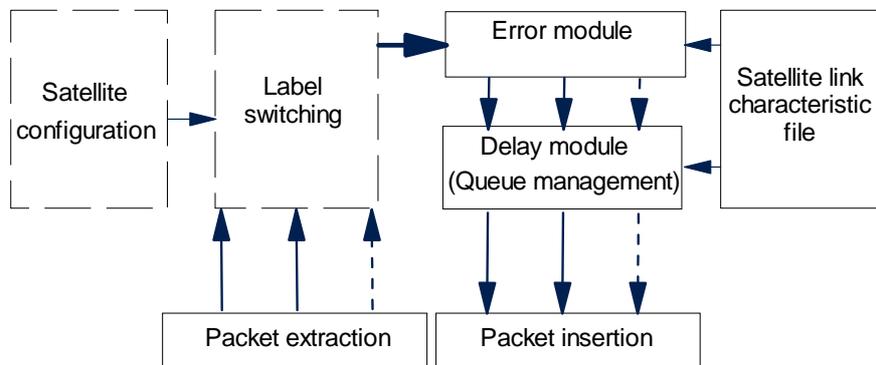


Figure 44 : Blocs Fonctionnels de l'émulateur satellite

Le module « Label Switching » est configuré par des tables de connectivité chacune associée à un transpondeur DVB-RCS et permettant de déterminer le spot de descente en fonction de leur label MAC. Il est également en mesure de répliquer les paquets sur différents spots pour le support d'un service multicast par exemple. Enfin la commutation peut également être désactivée pour l'émulation de satellite transparent.

Les profils de perte peuvent être configurés manuellement ou selon des lois de distribution fonction de la bande de fréquence et de la forme d'onde considérées ou correspondant à des conditions météorologiques favorables ou non. Enfin on peut également faire appel à un module de jeu de pertes basées sur des traces réelles.

III.4.4 Configuration de la plateforme

La plateforme d'expérimentation offre de nombreux paramètres de configuration qui nécessitent d'être calibrés avec soin et pertinence pour se rapprocher au mieux des configurations standards des satellites actuels.

Nous allons nous intéresser à la configuration de référence adoptée lors de notre campagne d'évaluation de l'architecture de QoS et en particulier à la couche physique. La part la plus importante de cette configuration réside dans la structure de la supertrame qui conditionne à la fois la granularité des allocations, le taux de rafraîchissement des allocations et des requêtes de capacité.

La supertrame est composée de 10 trames de durée de 50 ms. Chaque burst ATM contient une cellule ATM de 53 octets (cette taille est modifiable). Les CRs sont émises à chaque début de supertrame soit toutes les 500ms. Il en va de même de l'émission du TBTP. Les débits de transmission des STs sont variables mais on les fixe selon les catégories de ST considérés

correspondant respectivement à 144 kbps, 384 kbps, 1024 kbps et 2048 kbps. La décomposition de la supertrame dans la plateforme n'est que temporelle et non fréquentielle. Les STs se partagent ainsi une unique porteuse au sein d'un même spot.

En ce qui concerne le satellite le délai est fixé à 250ms avec une gigue égale à $\pm 1ms$ et le profil d'erreur est typique de conditions météorologiques favorables¹.

Les autres paramètres concernent la partie QoS de la plateforme. Leur calibration fera l'objet d'une section à part dans le chapitre 4 dans l'évaluation de performance de l'architecture de QoS.

III.5 QoS dynamique pour les applications non conscientes de la QoS sous-jacente

A ce stade, l'architecture de QoS SATIP6 ne repose que sur les informations contenues dans l'en-tête des paquets IP (soit le quintuplet {adresse source, port source, adresse destination, port destination, protocole de transport}, soit le champ DSCP) pour classifier les flux applicatifs. Ainsi le lien entre les flots et leurs besoins en terme de QoS n'est pas toujours évident puisque toutes les applications n'utilisent pas des ports réservés voir les négocient dynamiquement à l'établissement d'une session. Les filtres multi-champs basés sur des ports réservés ne sont malheureusement d'aucune utilité dans ces cas-là. De ce fait, il est relativement difficile de réaliser une classification pertinente ou un contrôle d'admission efficace si l'on n'est pas en mesure d'identifier l'application et les ressources nécessaires à son bon fonctionnement.

Notre avons donc doté cette architecture de deux solutions applicatives novatrices permettant de rétablir ce lien entre l'architecture de QoS et les applications. Dans la première approche, les règles du classificateur du ST sont dynamiquement modifiées à partir d'informations de niveau transport délivrées par un agent localisé dans les équipements utilisateurs. Ce dernier est à même de faire le lien entre une application et les différents flots la composant et donc de requérir, à l'initiative de l'utilisateur ou l'administrateur, un niveau de QoS pour l'application ou l'un de ces flux disponibles sur le segment satellite. La deuxième approche s'appuie sur le protocole d'établissement de session SIP qui véhicule, en plus des caractéristiques de niveau transport sur les flots, des informations de niveau applicatif tels que les codecs et le type de média impliqué indispensable pour un contrôle d'admission dynamique.

III.5.1 Un agent de sélection pour les applications non adaptées à la QoS

III.5.1.1 De la QoS pour toutes les applications

Très peu d'applications sont aujourd'hui programmées pour être « conscientes » des mécanismes de QoS offerts par les réseaux sous-jacents. Comme les réseaux satellites nécessitent une gestion fine de la QoS et que les applications sont très souvent incapables de spécifier leurs besoins, notre proposition d'agent de QoS offre une solution « orientée utilisateur » permettant à toutes les applications de bénéficier au mieux des services du réseau.

Installé sur le terminal utilisateur, l'agent de QoS détecte les flux applicatifs, les associe statiquement ou dynamiquement avec l'intervention de l'utilisateur au niveau de QoS désiré, et configure le terminal satellite (par l'intermédiaire du serveur de QoS) afin qu'il prenne en compte la demande.

La Figure 45 rappelle l'architecture du système satellite et le rôle du terminal satellite dans la classification des flux et leur émission selon la politique choisie. L'utilisateur peut donc, à distance, grâce à l'usage des agents de QoS (localisé dans son terminal) et du serveur (localisé dans le ST), « configurer » son terminal satellite pour obtenir le service qu'il souhaite appliquer à une de ses applications en cours d'utilisation.

¹ Cette notion subjective correspond, dans l'émulateur satellite, à un jeu de traces satellites réelles.

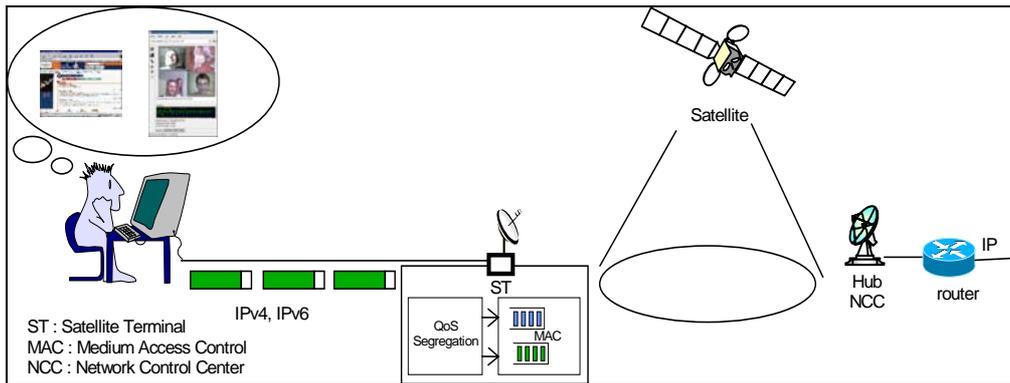


Figure 45 : Gestion de QoS par le ST

III.5.1.2 Principe de fonctionnement

L'agent de QoS a été conçu pour offrir un mécanisme de signalisation des besoins de QoS aux applications qui ne sont pas programmées en conséquence et ce sans aucune modification. Excepté le terminal utilisateur, aucune entité du réseau satellite (y compris son propre ST) ne peut avoir une vue précise des applications que l'utilisateur exécute. En effet, lorsqu'un paquet est émis sur le réseau, toute référence à l'application émettrice est masquée par les encapsulations successives. A part pour certaines applications dont les numéros de ports sont fixes et connus, il n'est pas possible pour une entité extérieure, à moins de désencapsuler systématiquement tous les paquets, d'identifier l'application émettrice et par conséquent la QoS qu'elle pourrait requérir. Cependant ce processus reste relativement lourd pour des terminaux utilisateurs en particulier sur des terminaux de petite envergure (e.g. PDA).

L'agent de QoS est donc un « démon » installé sur le terminal utilisateur qui détecte toutes les connexions sortantes des applications en cours d'utilisation. Une interface graphique (Figure 46) présente à l'utilisateur à tout moment l'état de ses connexions réseau et les applications qui en font usage.

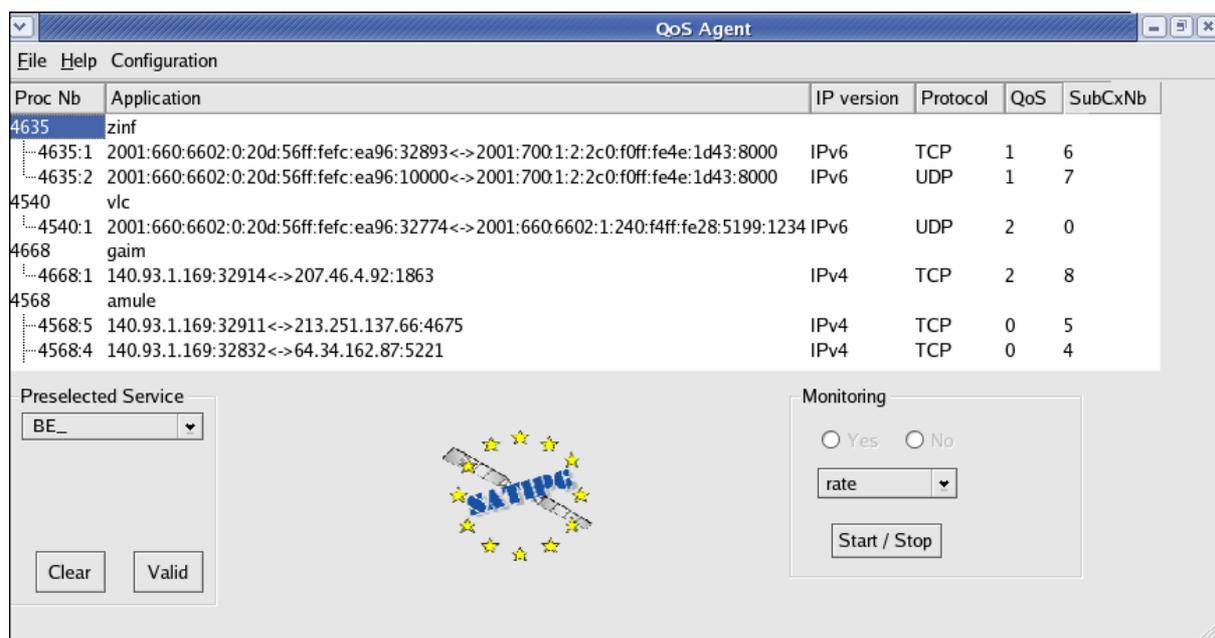


Figure 46 : Interface graphique de l'agent de QoS

Un module permet de sélectionner pour une application, c'est à dire pour l'ensemble des flux

qu'elle manipule, ou pour un seul flux, un des services disponibles et de signaler ce choix au terminal satellite. Chaque fois que l'utilisateur sélectionne un service pour une application donnée, l'agent de QoS signale au serveur de QoS situé sur le ST la liste des connexions concernées et le numéro du service associé par un protocole transactionnel dédié. Si les ressources ne sont pas disponibles, l'agent de QoS en est informé immédiatement. Avec les informations reçues, le serveur de QoS est capable de marquer et de rediriger les paquets issus du terminal utilisateur vers la file de niveau IP correspondant au service requis dont l'architecture décrite en section III.3 assure le bon fonctionnement.

Il est à noter que ce système n'agit que sur les connexions sortantes car notre architecture vise la gestion de QoS sur le lien DVB-RCS montant. Toutefois, dans le cas d'une communication bidirectionnelle et dans un scénario où les deux utilisateurs sont sur le réseau satellite, si le récepteur souhaite une amélioration d'un flux reçu, il doit la négocier avec l'hôte émetteur afin que celui-ci sélectionne (et paye en conséquence) un meilleur service.

III.5.1.3 Cas d'utilisations

Ce paragraphe détaille plus précisément deux cas réels dans lesquels l'agent de QoS a pu être utilisé sur la plateforme d'émulation.

III.5.1.3.1 Sélection d'un service « Voix » pour une application de VoIP

Ce premier cas considère une entreprise composée de plusieurs filiales qui offre l'utilisation de la voix sur IP afin de réaliser des économies de communications. Elle recommande d'autre part, « quand cela est possible », à ses utilisateurs d'utiliser un service moins cher que le service garanti nommé « voix » qui est plus taxé qu'une communication téléphonique classique.

Ainsi, les agents de QoS activés par défaut sur l'ensemble des postes de l'entreprise sont préconfigurés par le service informatique pour associer toutes les communications de voix sur IP (par exemple les applications Gnomemeeting, Microsoft Messenger, ...) au service AF. A l'initialisation d'une communication, les agents de QoS de l'appelant et de l'appelé configurent les STs pour appliquer le service approprié aux flux échangés.

Si, durant une conversation, la qualité du service AF n'est plus suffisante pour garantir une qualité de conversation suffisante, les utilisateurs ont la possibilité de changer de service grâce à l'agent de QoS. Pour cela, les utilisateurs ouvrent l'interface graphique et sélectionnent pour l'application de VoIP utilisée le service EF.

Dans ce cas, la qualité est garantie et permet de continuer la communication malgré un surcoût temporaire. A la fin de la communication, dès la coupure des connexions réseau, l'agent de QoS informe le terminal satellite qu'il n'est plus nécessaire de maintenir cette association de QoS. La taxation s'arrête dès cet instant.

III.5.1.3.2 Sélection de services de QoS pour des applications aux ports non réservés

Dans la Figure 46, nous avons 4 applications utilisant des ports non réservés et pour lesquelles l'administrateur ne peut créer des filtres statiques. Ces applications sont :

- Un serveur de streaming vidéo, VideoLanClient, référencé sous le nom de commande 'vlc'
- Un serveur de streaming MP3 zinf
- Une application de messagerie instantanée, Gaim
- Une application de Peer-to-Peer (P2P), Amule qui est la version Linux du logiciel de P2P Emule.

Ces applications comme nous pouvons le voir sur la Figure 46 n'utilisent pas de ports réservés. Des classes de service leur ont été attribuées allant de 0 à 2 : 0 pour le BE, 1 pour un service de type EF et 2 pour le service AF. Le streaming mp3 étant effectué à travers une connexion TCP et les temps de bufferisation étant relativement court, le service de QoS maximum est requis. Le streaming de vidéo s'effectuant en mode non connecté, un service de

type AF est considéré comme suffisant. Pour le P2P particulièrement gourmand en bande passante mais tolérant pertes et délais, la priorité minimale a été affectée afin de ne pas compromettre la QoS offerte aux autres applications.

III.5.1.4 L'implémentation

L'idée maîtresse lors du développement de l'agent de QoS était de disposer d'une application portable sous Windows et Linux, les systèmes d'exploitation considérés dans les terminaux utilisateurs. Notre choix s'est porté sur QT qui offre un éditeur d'interface graphique (QT Designer) et un environnement de programmation C++ multi-plateforme.

Le GUI de l'agent de QoS (cf. Figure 46) a entièrement été conçu grâce à QT Designer.

En plus du support des IHMs, QT met à disposition un ensemble de fonctionnalités portables à travers sa bibliothèque dont notamment :

- La gestion des threads (le serveur de QoS est un serveur multi-threadé)
- Des sockets à travers les *Q.Sockets* permettant une gestion simplifiée d'IPv6 et IPv4
- La génération et l'analyse de contenu XML (eXtensible Markup Language): (SAX (Simple API for XML) et DOM (Document Object Model))
- La manipulation et le traitement des chaînes de caractères

L'association entre les applications et leurs flots respectifs est effectuée par le post-traitement des sorties obtenues à partir des commandes « lsof » sous Linux et « netstat » sous Windows.

Le protocole de communication entre le serveur de QoS et l'agent de QoS est basé sur TCP et la structuration des messages échangés est au format XML. Le choix de XML s'est imposé du fait de la facilité d'évolution des attributs des messages, de sa portabilité (le SIP proxy détaillé en section III.5.2 dialogue également avec le serveur de QoS et du nombres conséquents d'outils XML dans tous les langages de programmation courants.

Le format des requêtes XML de l'agent de QoS est illustré Figure 47.

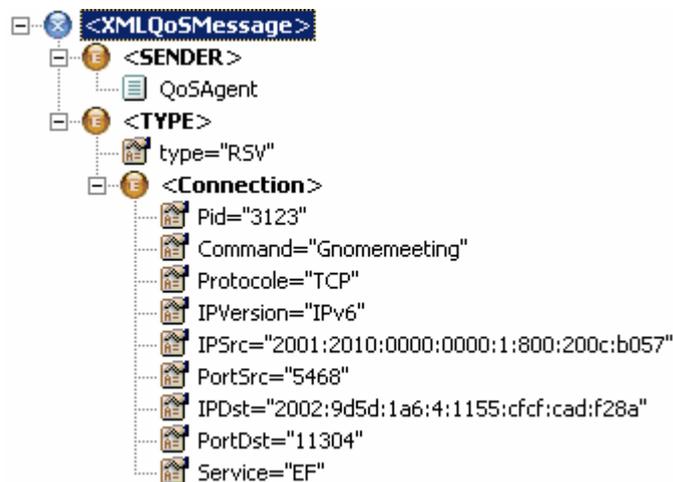


Figure 47 : Format XML véhiculé par l'agent de QoS

Le premier attribut d'une connexion est le PID de la commande sur le terminal utilisateur. Le nom de la commande correspond au nom de l'application qui peut permettre au ST d'appliquer une politique particulière en fonction des paramètres de la connexion et du nom de l'application si son comportement est connu de l'administrateur. Ensuite nous retrouvons le quintuplet <IP Src, IP Dst, Port Src, Port Dst, type de protocole> nécessaire à l'identification d'une connexion issue d'une application. Le dernier attribut est le champ Service lié aux classes de service DiffServ implémentées dans le ST. Une description plus intelligible des services pour l'utilisateur est envisageable.

Les diagrammes de classes suivants définissent la structure du module de communication de

l'agent de QoS Figure 48(a) et du module du Traceur de connexion Figure 48(b).

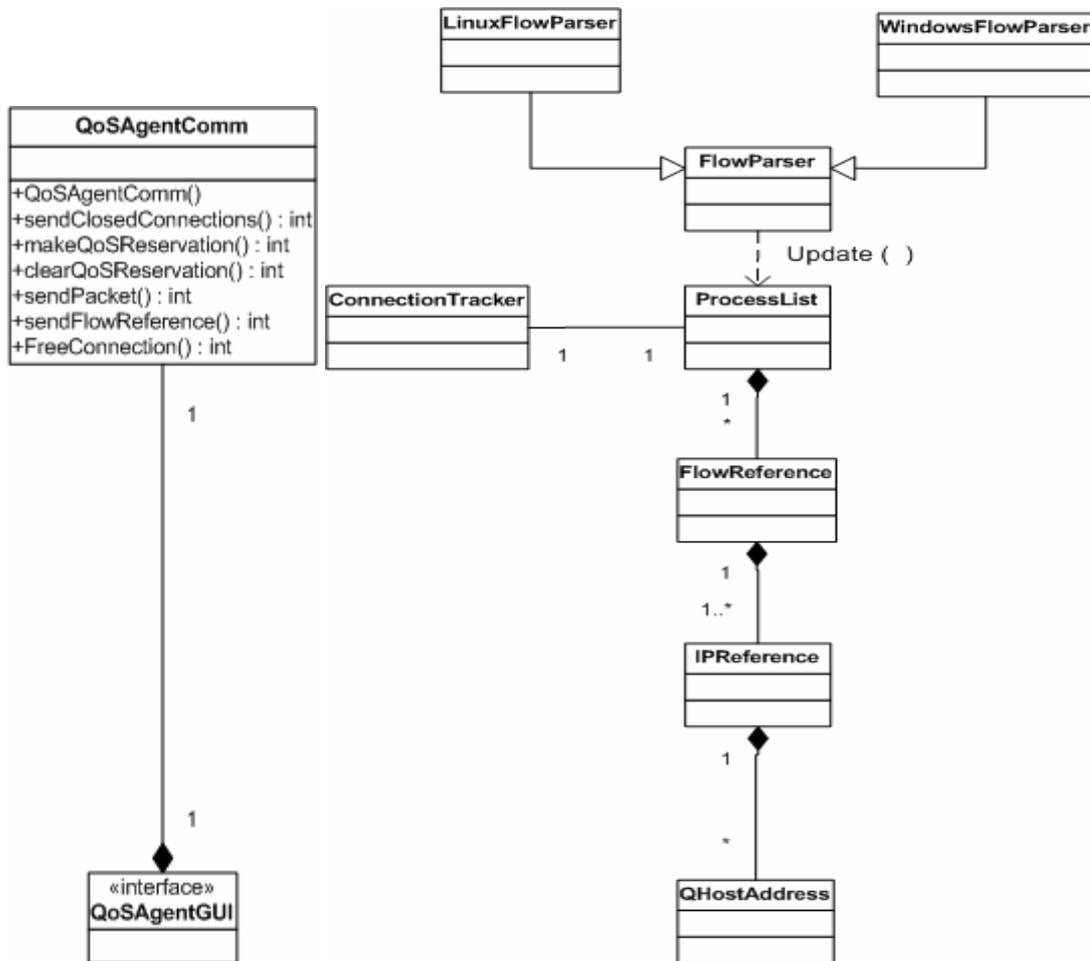


Figure 48 : Module de communication (a) et du Traceur de connexions (b) de l'agent de QoS

La classe QT *QHostAddress* permet de créer une instance d'une adresse IP qui est indépendante de la plateforme et du protocole utilisé. Elle permet de stocker des adresses IPv6 et IPv4 et d'y accéder simplement indépendamment du type de plateforme considéré. La classe *FlowReference* permet d'identifier un flot par les adresses et ports source et destination. La classe *ProcessList* consiste en un vecteur de *FlowReferences* et la classe *FlowParser* est responsable de la mise à jour de ce vecteur.

Enfin le diagramme de classe du *QoSServer* est donné dans la Figure 49. Le serveur de QoS crée une instance *ClientSocket* pour chaque nouvelle connexion qui gère la mise à jour du vecteur des réservations du serveur de QoS. Le classificateur du ST fait appel à la primitive *getAssignedQoS()* du serveur de QoS pour identifier le service associé à un flux. S'il n'apparaît pas dans la liste des réservations, le paquet est orienté par défaut dans le service BE.

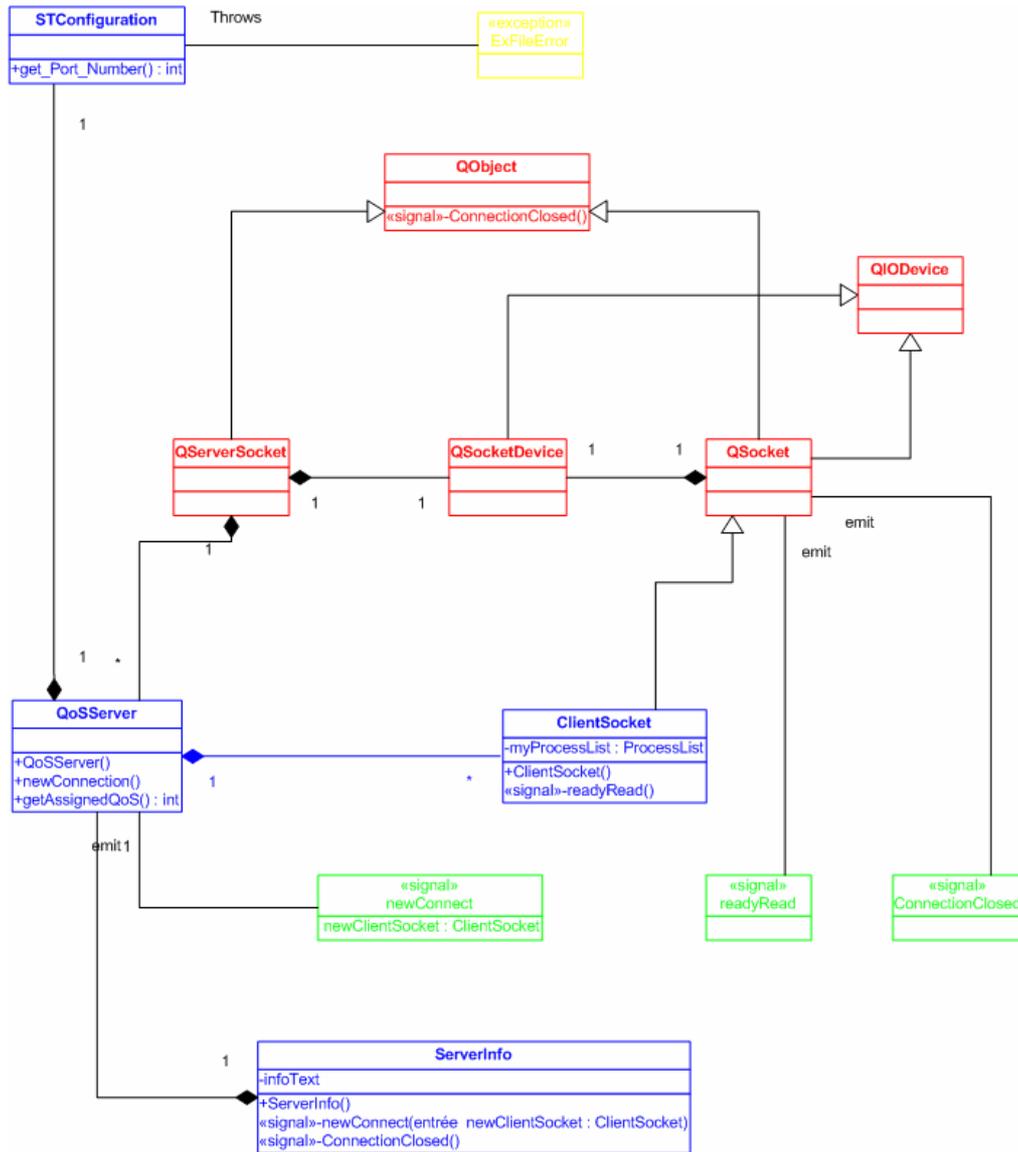


Figure 49 : Diagramme de classe du serveur de QoS

III.5.1.5 Les extensions envisagées

Le mode de fonctionnement de l'agent de QoS est ici résolument tourné vers l'utilisateur mais un mode « administrateur » de l'agent de QoS est également envisagé. Le but de ce démon serait de délocaliser des fonctionnalités du réseau dans le terminal client et de réaliser un certain nombre d'opérations de QoS comme le marquage DiffServ des flux applicatifs et le conditionnement du trafic directement dans le terminal utilisateur. Ces fonctionnalités seraient d'autant plus intéressantes sur un réseau sans fil local où l'agent de QoS en régulant les flux à partir du terminal utilisateur éviterait le gaspillage ou le vol de ressource dédiée aux applications temps réel. De plus, de nombreuses fonctionnalités de supervision notamment vis-à-vis de la QoS expérimentée par application permettraient de mieux dimensionner le réseau satellite en l'adaptant aux comportements des utilisateurs et de leurs applications. Afin de pouvoir réaliser ce scénario, il faudrait toutefois que l'agent de QoS soit relié à un serveur central qui mettrait à jour la politique de QoS à appliquer en fonction de la politique de l'administrateur du réseau satellite et du profil de l'utilisateur.

Plusieurs évolutions de l'agent de QoS sont d'ores et déjà prévues et sont particulièrement favorisées par la programmation Objet, l'utilisation d'un code source portable et d'un format de

structure de données basé sur XML. Parmi elles, nous noterons principalement le chargement dynamique de politique de QoS au démarrage de l'agent de QoS (actuellement statique), l'évolution du protocole de transport (SOAP, Compression de donnée, Sécurisation des échanges de message, Authentification, UDP pour les réseaux WiFi...), la possibilité de dialoguer directement entre agents de QoS et la constitution d'une entité centrale supervisant et configurant l'ensemble des serveurs de QoS et leurs modules de contrôle d'admission distribués dans les STs.

III.5.1.6 Les limites

L'agent de QoS à travers son mode de réservation par application offre une solution de réservation de ressource intuitive et simple pour les utilisateurs. Adapté pour des applications générant des flots applicatifs aux besoins de QoS homogènes, elle se révèle moins pertinente pour des applications multimédias nécessitant pour chaque flot des niveaux de QoS spécifiques.

Ainsi il est impossible pour l'agent de QoS de distinguer à partir des ports source et destination un flux Audio d'un flux Vidéo car il n'existe pas de ports dédiés au transport de données vidéo ou audio.

Dans ce cas de figure, la seule solution est de déduire ces informations directement au niveau applicatif ou à partir des descriptions de médias véhiculées dans les protocoles d'établissement de session. Cette solution, détaillée dans la section suivante, a été spécifiée et implémentée à part du projet SATIP6.

III.5.2 Signalisation de QoS par un Proxy SIP

Une solution alternative à celle de l'agent de QoS est donc d'automatiser la réservation de ressource et ainsi de la rendre transparente aux utilisateurs qui ne seraient pas en mesure de choisir les classes de service les plus appropriées aux différents flux (audio et vidéo par exemple) générés par leur application. L'utilisation d'un protocole d'initiation de session véhiculant les caractéristiques de la session peut servir de support au déclenchement de cette réservation. Le protocole que nous avons retenu du fait de son succès actuel et du nombre croissant d'applications l'utilisant aussi bien dans le domaine public (code source libre) que privé (entreprise), est le protocole SIP.

Ce deuxième mécanisme permet la mise en place de la QoS de manière transparente grâce à l'interprétation des messages SIP échangés entre les applications au niveau d'une entité intermédiaire, le proxy SIP. De ce fait, l'application n'a pas à subir de modification. La seule condition à respecter est que ces applications soient compatibles avec les spécifications SIP suivantes [10][11].

III.5.2.1 Un proxy SIP amélioré

L'entité à laquelle la réservation de QoS est déléguée est un proxy SIP standard auquel nous ajoutons des fonctionnalités avancées. La fonctionnalité première d'un proxy SIP est de relayer les messages SIP d'entités SIP en entités SIP jusqu'au destinataire. Afin de rendre la réservation de QoS transparente vis-à-vis de l'application, le proxy SIP amélioré intercepte les descriptions de session véhiculées par SIP entre l'appelant et l'appelé, en déduit les caractéristiques de chaque média impliqué dans la session et effectue les réservations et les libérations de ressource associées auprès du serveur de QoS au nom de l'application.

Pour répondre à nos exigences, ce proxy doit fonctionner de manière « stateful » afin de maintenir l'état de la session et d'assurer la libération des ressources. Pour cela, toutes les requêtes et réponses SIP doivent passer par ce même proxy, ce qui est possible si cela est signalé dans le champ Record-Route de l'entête SIP d'un message véhiculant un INVITE. Par ailleurs, le protocole de transport retenu est UDP plus adapté que TCP classiquement exploité sur un lien satellite. En effet, le mode d'établissement de connexions nécessaires UDP nous oblige également à adapter la durée des timers de retransmission généralement calqué sur un RTT maximum de

500 ms ce qui est le minimum pour un aller/retour sur le réseau satellite.

Le proxy SIP se voit enrichi des fonctionnalités suivantes :

- un *analyseur SDP* rendant le proxy à même d'interpréter les descriptions de session répondant au format SDP véhiculées dans les messages SIP.
- Une *table des médias* qui est mise à jour durant l'établissement des sessions SIP. Chaque média est caractérisé par le quadruplet (IP Src, IP Dst, Port Src, Port Dst), le type de média (Audio, Vidéo ...) et le profil RTP associé. Les médias négociés entre l'appelant et l'appelé sont identifiés par l'intermédiaire d'un Call-ID, identifiant unique de la session SIP concernée.
- Un *mapping SDP/DiffServ* qui assure la correspondance entre le type de média (et le codec) et le service DiffServ (et le débit crête associé), conversion indispensable au module de QoS.
- Un *module de QoS* qui prend en charge la réservation des ressources associées à chaque média d'une session SIP auprès du serveur de QoS.

En ce qui concerne le traitement prioritaire de la signalisation SIP, la centralisation de la signalisation SIP par un proxy local permet également de simplifier sa classification dans le ST puisque tous les messages SIP sont émis par la même adresse source qui est celle du proxy.

III.5.2.2 Localisation du proxy

La réservation de QoS pour le trafic empruntant la voie montante du satellite a toujours lieu auprès du ST qui en a la charge. Un proxy SIP avec QoS est déployé dans chacun des LANs interconnectés par le satellite. Cette architecture décentralisée se révèle particulièrement adaptée :

- aux exigences de passage à l'échelle en repoussant la gestion de la QoS par flux dans les LANs d'extrémité du réseau satellite
- au respect des délais d'établissement de session en minimisant le nombre d'allers-retours de la signalisation de session et de QoS sur le lien satellite. On considère ainsi, selon [119], un temps d'établissement de 1.5 s comme étant un impératif pour la meilleure classe de trafic tandis que pour le Best-Effort, on peut se contenter de 7 s.

III.5.2.3 Cas d'utilisation

Les Figure 50 et Figure 51 illustrent le provisionnement dynamique des ressources à l'établissement d'une session où deux interlocuteurs souhaitent communiquer par l'intermédiaire d'une visioconférence. Le proxy intercepte et convertit les descriptions SDP pour chaque média (Audio G711 et Vidéo H263), puis il effectue les réservations de ressource correspondantes auprès du serveur de QoS. Les modes de réservation décrits ici sont le mode « QoS Enabled » et QoS Assured. Les modes de réservation sont détaillés dans III.5.2.4.

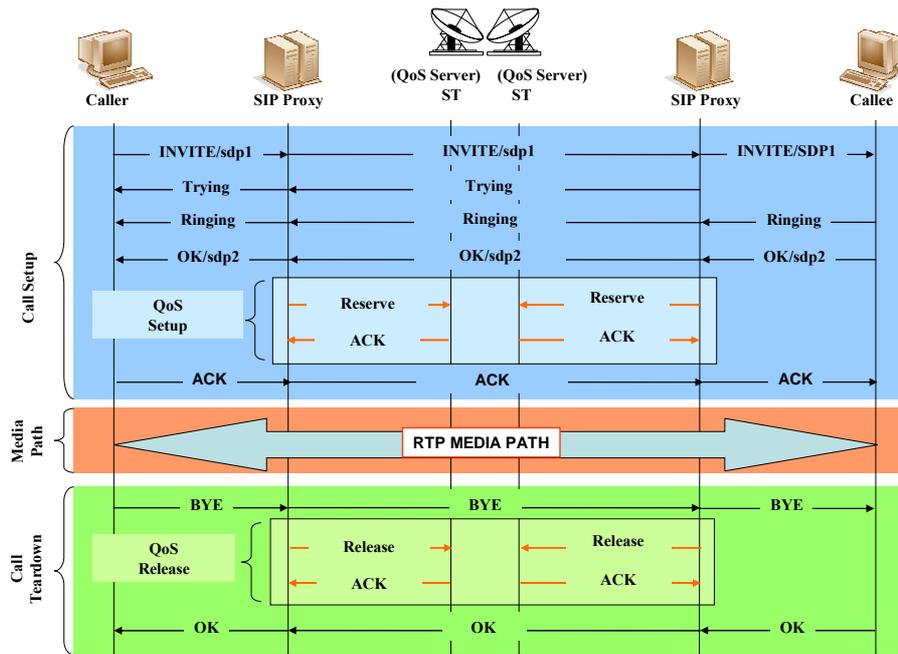


Figure 50 : Ssion SIP avec réservation de QoS par le proxy en mode « QoS Enabled »

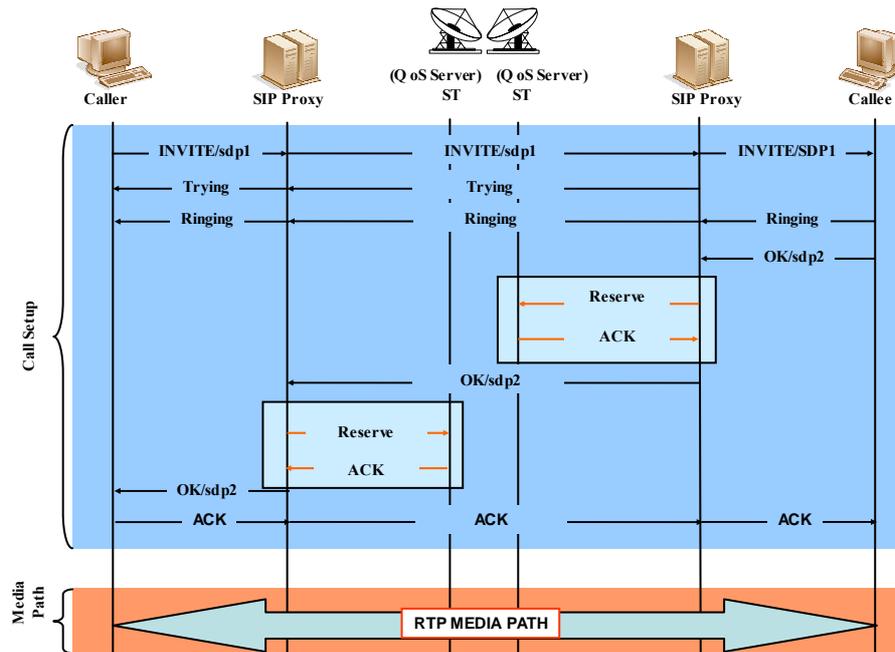


Figure 51 : Session SIP avec réservation de QoS par le proxy en mode « QoS Assured »

III.5.2.4 Mode de réservation : QoS Enabled ou QoS Assured

Deux modes de réservation sont supportés par le proxy : QoS Enabled et QoS Assured. Le premier mode permet l'établissement de la session entre les deux équipements terminaux avec ou sans la mise en place des ressources nécessaires. Le deuxième mode impose la réservation de ressource comme une précondition à l'établissement de la session multimédia.

Les services de téléphonie afin qu'ils soient adoptés par le plus grand nombre doivent offrir des garanties de QoS comparables à ceux du téléphone sur RTC. Le mode « QoS Assured »

offre cette possibilité en accroissant toutefois le temps d'établissement de la session et en imposant le support de méthodes SIP supplémentaires (UPDATE, PRACK et SESSION PROGRESS) pas toujours implémentées dans les piles SIP actuelles ou prises en compte par les applications. Il augmente ainsi sensiblement le délai « post-dial »¹ mais réduit le délai « post-pickup »² ainsi que les probabilités d'échec d'un appel ayant abouti du fait du manque de ressources pour la communication. Un mode « QoS Assured » sur une session basique est toutefois implémenté en forçant le proxy à attendre la confirmation d'une réservation de ressource avant de relayer le message OK. Le mode QoS Enabled réduit quant à lui le délai « post-dial » et reste compatible avec le maximum d'applications existantes mais il peut entraîner des coupures ou distorsions de la voix au début de la communication voire la mise en place d'une communication de très mauvaise qualité si elle s'appuie sur le service Best-Effort. Cependant ce mode offre un bon compromis entre Best Effort et le mode « QoS Assured » en offrant un temps d'établissement de session réduit et une qualité de service accrue si les ressources sont disponibles.

III.5.2.5 L'implémentation

L'ensemble des fonctionnalités avancées décrites en section III.5.2.1 ont été ajoutées à un proxy existant, le NIST-SIP-Proxy [120], basé sur une pile SIP Java respectant la spécification JAIN-SIP 1.1 et dont le code source était disponible. Les fonctionnalités ont été implémentées afin que le minimum de changement soit nécessaire dans le code existant et qu'elles soient le plus modulaire possible. Le choix de Java nous a permis également de tester le proxy sous Windows et Linux.

Le protocole adopté pour communiquer avec le serveur de QoS est également basé sur XML et communique les informations sous un format proche de celui communiqué par l'agent de QoS. Dans la Figure 52, le proxy SIP effectue une réservation de QoS auprès du serveur de QoS pour une session de type VoIP basée sur le codec G711 dont le débit crête est fixé à 75 kbps au niveau IP.



Figure 52 : Format XML communiqué par le proxy SIP

III.5.2.6 Les extensions

Alternativement au mode distribué du contrôle de ressource (Serveur de QoS dans chaque

¹ Le délai « post-dial » correspond dans le cas d'une session SIP à l'intervalle de temps séparant l'envoi de l'INVITE et la réponse RINGING l'avertissant de la sonnerie du téléphone de son correspondant.

² Le délai « post-pickup » correspond au délai de mise en place de la connexion pour les communications après que le correspondant ait décroché.

ST), on envisage un gestionnaire de ressource central localisé derrière le NCC qui répondrait aux différentes réservations de ressources de proxys SIP améliorés ou délèguerait ce rôle au serveur de QoS dans les STs. Les politiques de « mapping » des SDPs en paramètres de QoS seraient alors directement chargées par les serveurs de QoS à partir de ce serveur central ainsi que les algorithmes de contrôle d'admission.

Une autre extension consisterait à distribuer le proxy SIP amélioré dans chaque terminal utilisateur. Il constituerait une extension du réseau qui adapterait, de manière transparente pour l'utilisateur, le mode de réservation de ressource en fonction du type d'applications SIP utilisées et de SDPs véhiculés. Il pourra également piloter la configuration d'un module de conditionnement et de marquage DiffServ du trafic en sortie du terminal utilisateur.

Enfin les mêmes extensions que pour l'agent de QoS sont envisagées en terme d'amélioration du protocole de communication avec le serveur de QoS.

III.6 Conclusion

Dans ce chapitre, nous avons spécifié une architecture de QoS avancée pour les systèmes DVB-RCS/S. Ces systèmes vont être amenés dans un futur proche à supporter des types de trafic IP fortement hétérogènes et la gestion statique actuelle des ressources est largement sous-optimale par rapport aux profils sporadiques et l'interactivité des services large bande multimédias. De plus, ces mécanismes souvent propriétaires ne sont pas interopérables avec les mécanismes d'administration de ressource orientés IP largement plus développés dans les réseaux terrestres.

De ce fait, l'architecture proposée a tout d'abord défini un mécanisme d'allocation de bande passante à la demande au niveau 2 complet afin de partager au mieux les ressources satellites en fonction des besoins ponctuels en bande passante mais aussi en QoS de chaque ST.

Une architecture DiffServ a ensuite été « mappée » sur cette dernière pour trois raisons principales : le passage à l'échelle, l'interopérabilité avec les futures architectures NGN des réseaux terrestres basées sur IP et la possibilité pour les applications de bénéficier de services différenciés standardisés de niveau IP interagissant, étroitement et de manière transparente, avec les mécanismes de QoS de niveau 2.

Enfin, partant du constat que la majorité des applications courantes multimédias n'intègrent pas de protocoles de réservation de ressources, nous avons mis en œuvre deux solutions applicatives permettant à ces applications de bénéficier des services différenciés implémentés sur le segment satellite. La première solution est basée sur un « agent de QoS » implémenté dans le terminal utilisateur et permettant à l'utilisateur d'associer dynamiquement un service de QoS à l'ensemble des connexions sortantes de n'importe quelle application. La deuxième solution s'appuie sur la délégation de la réservation de ressource à une entité intermédiaire, un proxy SIP, capable d'extraire les caractéristiques de chaque média négocié au cours de l'établissement d'une session multimédia et de réserver les ressources correspondantes en conséquence.

Afin de démontrer la pertinence de notre architecture, elle a été implémentée dans une plateforme d'expérimentation émulant un réseau satellite de type DVB-S/RCS. Les résultats de cette évaluation font l'objet du chapitre 4 où les scénarios de mesure ainsi que les outils et les métriques utilisées y sont consignés.

Chapitre 4 : Evaluation de l'Architecture de QoS DVB-RCS

IV.1 Introduction

Afin de valider l'architecture de QoS DVB-RCS que nous avons définie et qui est implémentée dans SATIP6, une campagne de tests a été menée. Nous insistons particulièrement sur l'évaluation des performances de QoS applicatives perçues de bout en bout sur la plate-forme d'émulation.

Cette évaluation porte principalement sur les applications multimédias de type voix sur IP (VoIP), visioconférence et vidéo à la demande (VoD). Ce choix s'explique par la part de plus en plus importante qu'elles vont jouer dans les NGNs. Fortement consommatrices en bande passante, leur interactivité et leur profil de trafic mettent également à rude épreuve les architectures de QoS.

La première partie s'attache à caractériser le comportement de ces applications à partir de traces issues de sessions multimédias réelles. Nous déduisons de cette analyse la difficulté de caractériser par un modèle déterministe ou statistique une source de trafic multimédia. Nous démontrons ainsi l'intérêt d'un outil de rejeu de traces réelles afin de reproduire fidèlement le comportement de trafics multimédias sur la plate-forme d'émulation. Ensuite, nous décrivons l'outil de rejeu que nous avons développé pour les besoins de cette évaluation. Nous détaillons également la méthode et les outils d'analyse utilisés pour évaluer les performances de l'architecture.

Dans une deuxième partie, nous proposons une calibration de la plate-forme, c'est-à-dire un réglage des différents mécanismes de QoS de l'architecture en adéquation avec les caractéristiques du trafic multimédia observé précédemment.

La troisième partie décrit les résultats de la campagne de test que nous avons effectué sur la plate-forme SATIP6. Plusieurs scénarios ainsi que différentes configurations de la plate-forme ont été envisagés.

Enfin la dernière partie présente les conclusions tirées de ces expérimentations et les améliorations possibles.

IV.2 Caractérisation des applications multimédias

IV.2.1 L'enjeu

Les enjeux d'une telle caractérisation sont multiples. Les applications multimédias génèrent du trafic aux caractéristiques singulièrement différentes des applications de données classiques. Leur bon fonctionnement est contraint par le respect de délais, gigue et taux de pertes stricts. Par exemple, la qualité d'une communication vocale est fortement sensible aux pertes et aux délais et un trafic vidéo peut être significativement dégradé lors des périodes de congestion du réseau. L'analyse de traces capturées lors de sessions multimédias réelles révèle qu'un nombre restreint de critères suffit pour isoler des catégories d'applications.

Ces différents profils de trafic devant coexister au sein d'un même réseau, le support de services différenciés est indispensable. Cependant, au vu du nombre limité de services disponibles dans l'architecture SATIP6, il nous faut déterminer le degré de séparation nécessaire entre ces trafics c'est-à-dire : quels trafics doivent être séparés et quels sont ceux qui peuvent être mélangés au sein d'un même service ? La caractérisation joue un rôle essentiel pour répondre à cette question.

Enfin cette caractérisation est également indispensable pour dimensionner pertinemment l'ensemble des algorithmes de QoS liés au contrôle d'admission, au conditionnement et à l'ordonnancement des services EF et AF sur la voie retour. Elle nous permettra par la suite de déduire des informations véhiculées par les protocoles de session classiques les véritables besoins

en QoS (bande passante, type de média ...) des différents médias d'une même session.

IV.2.2 Recensement des applications multimédias

Les trois catégories, actuellement les plus répandues, à considérer sont les suivantes : la VoIP, la Visioconférence et la VoD. Un premier aperçu des contraintes de QoS de ces trois catégories d'applications est donné dans le Tableau 4.

Tableau 4 : Délai, Gigue et Pertes pour les applications audio et vidéo selon la recommandation Y.1540 ITU-T

Medium	Application	Degree of symmetry	Typical data rates	Key performance parameters and target values			
				One-way delay	Delay variation	Loss	Other
Audio	Conversational voice	Two-way	4 kb/s to 64 kb/s	< 150 ms preferred < 400 ms limit	< 1 ms	< 3 % Packet Loss Ratio (PLR)	
Audio	Voice messaging	Primarily one-way	4 kb/s to 32 kb/s	< 1 s for playback < 2 s for record	< 1 ms	< 3 % PLR	
Audio	High quality streaming audio	Primarily one-way	16 kb/s to 128 kb/s	< 10 s	< 1 ms	< 1 % PLR	
Video	Videophone	Two-way	16 kb/s to 384 kb/s	< 150 ms preferred < 400 ms limit		< 1 % PLR	Lip-synch: < 80 ms
Video	One-way	One-way	16 kb/s to 384 kb/s	< 10 s		< 1 % PLR	

IV.2.2.1 La Voix sur IP

IV.2.2.1.1 Caractérisation théorique

Dans le cadre de SATIP6, nous nous sommes principalement intéressés aux applications de Voix (IPv6 et IPv4) et au protocole d'établissement de session (SIP ou H323). Parmi elles, nous pouvons retenir : Linphone [122], Kphone [123], SIP Communicator [124], Gnomemeeting [125] et Windows Messenger [126] dans sa version 4.7.

L'ensemble de ces applications propose des communications de type audioconférence basées sur un codec choisi parmi une liste. Le choix du codec s'effectue en général automatiquement par l'application ou indirectement par l'utilisateur. Les principaux sont répertoriés dans le Tableau 5.

Tableau 5 : Caractéristiques des codecs VoIP

Codec	GSM 06.10 ¹	G711 ²	G.723.1 ³	iLBC ⁴
Taux d'échantillonnage	8Khz	8Khz	8Khz	8Khz
Débit du codeur	13kbps	64kbps	5.3kbps/6.3kbps	13.33kbps/15.2kbps
Période de paquetisation	20ms	30ms	30ms, 60ms, 90ms	30/20ms
Algorithme d'encodage	RPE-LTP	PCM (μ Law, ALaw)	CELP	LPC

Les applications d'audioconférence incluent un module de suppression de silence qui peut être activé ou non. Après la phase d'acquisition audio, le logiciel détecte les périodes de silence et

¹GSM 06.10 : [ETSI Recommendation GSM 06.10](http://kbs.cs.tu-berlin.de/~jutta/toast.html) & <http://kbs.cs.tu-berlin.de/~jutta/toast.html>

²ITU recommendation G711

³ITU Recommendation G.723

⁴iLBC : <http://www.ilbcfreeware.org/>

d'activité afin de ne transmettre des données que quand la source est active. Les périodes d'activité d'une source lors d'une communication téléphonique constituant de 30% à 60% du temps total. Le gain de bande passante est non négligeable.

Le signal est ensuite compressé et les échantillons audio codés sont collectés et stockés au sein d'un buffer. Enfin, ces échantillons sont encapsulés dans un paquet RTP destiné à être transmis sur le réseau à des intervalles temporels généralement fixes appelés « intervalle de paquetisation ».

Le surcoût protocolaire lié à cette encapsulation est, en théorie, entièrement défini, pour un codeur donné, par le couple (débit du codec, intervalle de paquetisation). Le Tableau 6 donne un aperçu des débits au niveau IP en fonction des codecs et de la taille des paquets RTP générés.

Tableau 6 : Débit IP par type de codeur

Codec	Débit (kbps)	Période de paquetisation (ms)	Taille de la charge utile RTP (octet)	Débit IP théorique (kbps)
G711	64	40	240	75
GSM	13.2	80	132	17
G723.1	5.3	30	20	16
G723.1	6.3	30	24	17
iLBC	13.3	20	50	24
iLBC	15.2	30	38	31

Afin de vérifier la validité de telles données théoriques, nous allons les confronter à des résultats empiriques obtenus à partir d'applications réelles.

IV.2.2.1.2 Caractérisation empirique

Avant l'analyse proprement dite des traces, nous allons décrire leur origine ainsi que leur degré de précision.

Ces traces ont été capturées par l'intermédiaire de l'outil tcpdump [127] à la sortie de l'interface réseau d'une machine hébergeant les applications multimédias. Par son intermédiaire, nous avons accès aux en-têtes des différents paquets émis par la machine ainsi qu'à leur temps d'émission. Des filtres de capture permettent d'isoler chaque connexion RTP séparément et d'en extraire les informations qui nous semblent les plus pertinentes. En l'occurrence, nous avons retenu la taille de la charge utile de niveau transport de chaque paquet et leur temps d'émission.

Les profils G711, GSM et SIREN générés par les applications Gnomemeeting et Windows Messenger ont ainsi été capturés. Le codec SIREN, quant à lui, utilisé pour la voix dans Windows Messenger est une extension du codec standard G722.1¹.

La Figure 53 illustre les profils de débit obtenus pour chaque application. Nous remarquons que, dans chaque cas de figure, le module de suppression de silence est activé. Les débits observés sont les débits utiles au niveau Transport. Sans la suppression de silence, on observe des débits constants durant l'ensemble de la session

¹ [ITU recommandation G722.1](#)

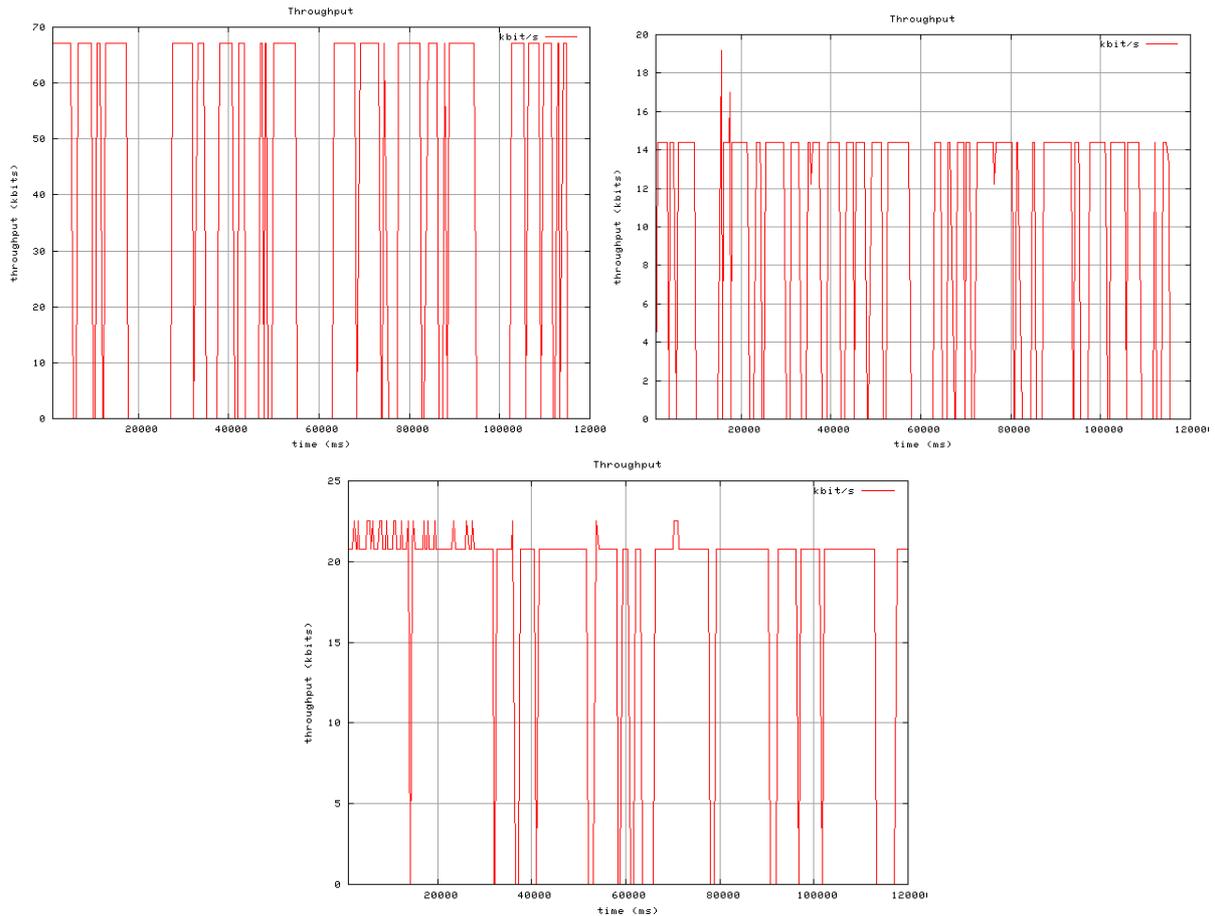


Figure 53 : Profil de débit d'audioconférence G711, GSM (Gnomemeeting) et SIREN (Windows Messenger) avec suppression de silence

Ces courbes ne nous renseignent que partiellement sur les propriétés intrinsèques de ce type de session et ne permettent de valider que l'hypothèse de débit constant avancée dans la section IV.2.2.1.1.

Pour le compléter, nous allons dans un premier temps déterminer la distribution de la taille des paquets et des délais « inter-paquet à la source » au sein des flux audio étudiés. La notion de délai inter-paquet à la source correspond à l'intervalle de temps séparant deux paquets consécutifs issu d'une même connexion à la sortie de l'interface réseau du terminal émetteur.

Quel que soit le codec utilisé, la taille des paquets dans une session audio est invariante (Charge utile RTP : GSM06.10 (132 octets), G711 (240 octets), SIREN (52 octets)) ce qui justifie bien l'hypothèse d'une paquetsation périodique à débit constant.

La Figure 54 illustre la distribution des délais inter-paquet à la source de paquet. Nous remarquerons que les délais inter-paquet à la source ne correspondent pas complètement aux intervalles de paquetsation théorique noté dans le Tableau 6. Ces différences peuvent s'expliquer par la charge éventuelle des terminaux utilisateurs, par la spécificité des applications utilisées. Ces différences constituent autant de paramètres qu'il est difficile de prendre en compte dans un modèle de trafic théorique. Ces détails ont également leur importance lors de la calibration des services de QoS de la plate-forme. Cet aspect est plus précisément évoqué dans la section IV.3.

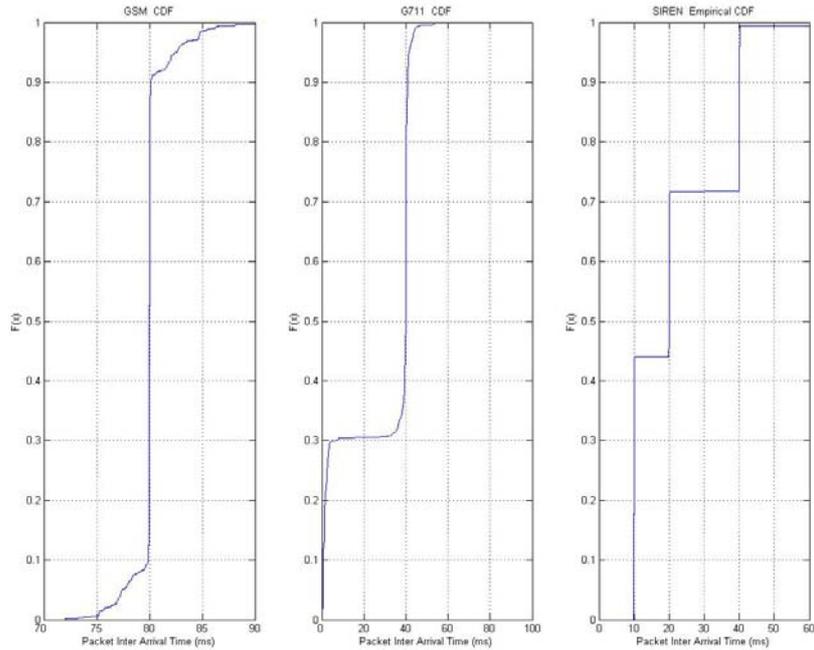


Figure 54 : CDF des délais inter-paquet à la source

IV.2.2.2 La vidéoconférence

Les sections suivantes abordent les sources de trafic vidéo. Ces dernières ont été découplées en deux catégories d'applications intégrant des techniques de compression de vidéo répondant à des impératifs et visant des objectifs différents. Le Tableau 7 répertorie notamment les débits, les formats vidéo et les techniques de compression généralement associés à des catégories d'applications.

Tableau 7 : Débits associés à différents services audiovisuels

	Video Format	Bit Rate required before compression	Bit Rate required after compression	Compression Standard
Mobile Videotelephony	SQCIF	4.4 Mbps	5 to 64 Kbps	MPEG4, H.263
VideoTelephony	QCIF	9.115 Mbps	32 to 64 Kbps	H.261, H.263
Videoconferencing	CIF	36.45 Mbps	112 Kbps	H.261
VCR-quality (DVI, CDI)	SIF	36.45 Mbps	1.5 Mbps	MPEG1
Studio-quality digital TV (DVD)	ITU-R BT.601	166 Mbps	3 to 6 Mbps	MPEG2
Digital High-Definition TV	SMPTE 240M	506 Mbps	25 to 35 Mbps	MPEG2
Super HD Video quality	SHD	> 6 Gbps	> 100 Mbps	MPEG2

Les sessions de visioconférence sont constituées de flux audio et vidéo non multiplexés et véhiculés à travers des connexions RTP indépendantes donc toutes les conclusions de la section précédente sont applicables à la partie audio d'une visioconférence.

Afin de réduire pertinemment le spectre de notre analyse, cette dernière est centrée sur la norme de codage H263¹ qui est la norme de compression vidéo la plus utilisée pour la visioconférence. Son algorithme de compression ressemble dans les grandes lignes à celui de la norme H261 mais en améliore significativement la correction d'erreur. Les améliorations en terme de prédiction et de codage permettent d'atteindre, pour une même qualité d'image, la

¹Recommandation ITU disponible <http://www.itu.int/rec/recommendation.asp?type=products&parent=T-REC-h>

moitié du débit H261.

L'encapsulation du flux H263 se fait également dans des paquets RTP. Les détails de cette encapsulation sont entièrement définis par la RFC 2190.

Les captures des sessions vidéo suivantes (Figure 55) ont été effectuées à partir de Gnomemeeting et de Windows Messenger

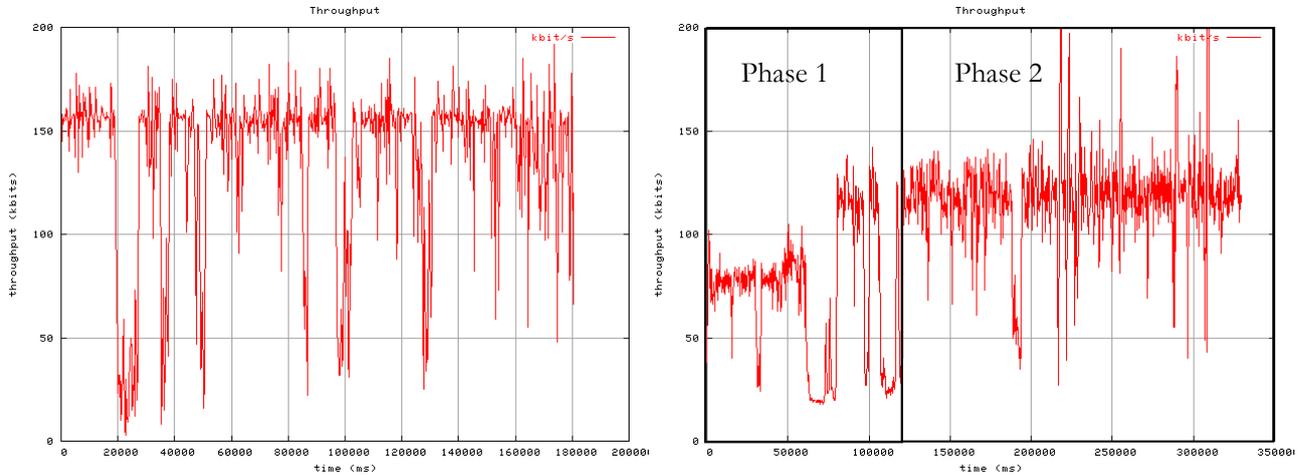


Figure 55 : Profil de débit de vidéoconférence sous Gnomemeeting (H261) (a) et Windows Messenger (H263) (b)

Dans le cas de Windows Messenger (à droite), les possibilités de configuration du codeur vidéo sont inexistantes tandis que pour Gnomemeeting, l'interface illustrée en Figure 56 permet de définir la bande passante (20KBps), le minimum de qualité requis par l'utilisateur (50%), le nombre d'images par secondes (4/s) et le nombre de blocs élémentaires constituant chaque image (4).

La session Windows Messenger est découpée en deux phases, une première où l'utilisateur derrière la WebCam est quasi immobile, la deuxième où l'utilisateur est en mouvement.

Une simple comparaison entre l'évolution des débits lors de ces deux phases confirme l'hypothèse d'un profil de trafic fortement dépendant du phénomène vidéo capturé.

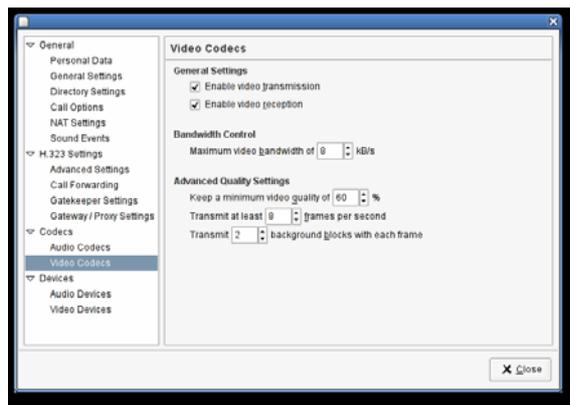


Figure 56 : Configuration du codeur vidéo H261 de Gnomemeeting

La Figure 57 représente la fonction de masse et de répartition empirique de la taille des paquets dans un flux vidéo généré par une application de type visioconférence et la Figure 58 illustre la distribution cumulative des délais inter-paquet à la source dans un flux Video.

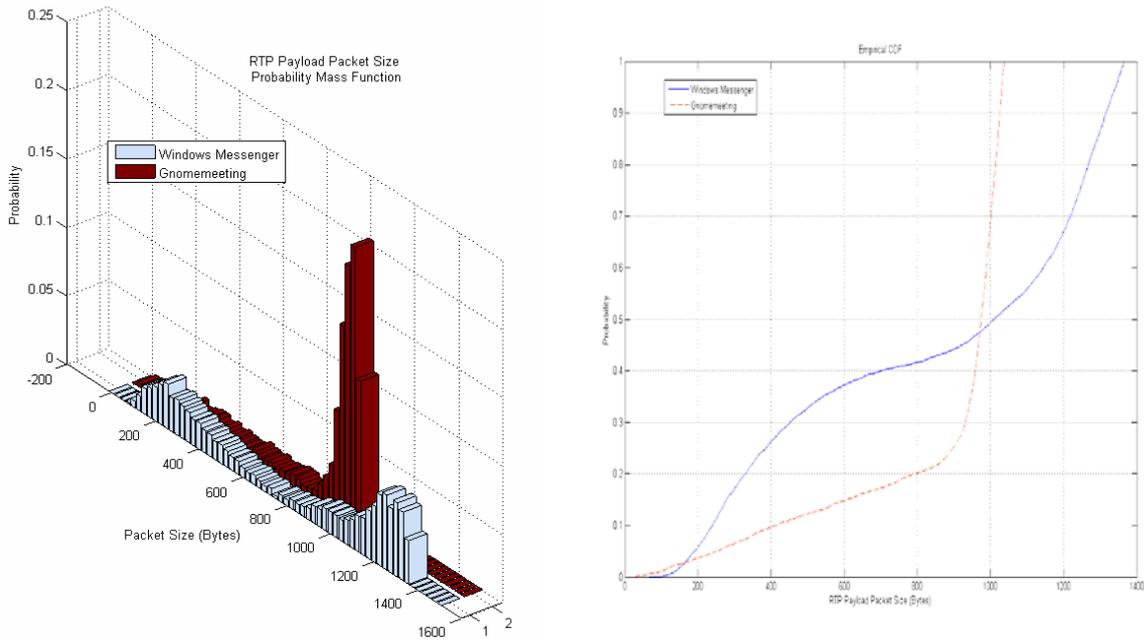


Figure 57 : Fonctions empiriques de répartition (a) et de distribution cumulative (b) de taille de paquets

Ces observations présentent des caractéristiques à partir desquelles il est difficile d'extrapoler un comportement générique. Ces profils sont dépendants du phénomène observé, de la configuration du codec (rafraîchissement, bande passante, qualité ...) et du type d'encapsulation des trames vidéo dans les paquets RTP.

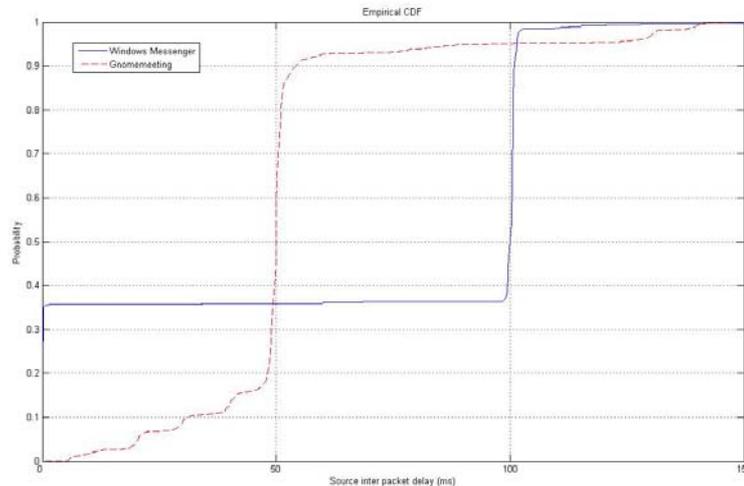


Figure 58 : CDF des délais inter-paquet à la source (vidéoConférence)

Les principales conclusions tirées de ces captures sont principalement la variabilité du débit et de la taille des paquets générés qui est elle-même fortement dépendante du type de phénomène observé à travers l'interface de capture vidéo (Webcam). Une étude approfondie des caractéristiques de ces connexions [128] nous montrent que contrairement à une source de voix, les sources de vidéo présentent des propriétés d'autosimilarité et de dépendance à long terme qui les rendent difficilement modélisables par des processus markoviens ou poissonniens.

Ces constatations démontrent la difficulté de simuler de tels trafics et donc la nécessité de recourir à des traces réelles afin de reproduire fidèlement le comportement d'une source de trafic vidéo.

IV.2.2.3 La vidéo à la demande

Les serveurs de diffusion vidéo enfin constituent la dernière catégorie d'application sur laquelle porte notre étude. Les applications de streaming, que ce soit pour la diffusion d'événements en direct ou la diffusion de média à la demande présentent des contraintes temporelles plus souples que la téléphonie ou la visioconférence. Ainsi l'interactivité n'apparaît pas toujours comme le critère d'évaluation de la QoS offerte exception faite des fonctions de contrôle de type magnétoscope. Les attentes des utilisateurs et les scénarios d'utilisation sont dès lors multiples, ne serait-ce qu'en terme de format et de qualité de vidéo :

- retransmission en direct ou en différé d'événement en qualité modeste (Microsoft Advanced Streaming Format, RealVideo, Quicktime) impliquant de faibles débits allant de quelques dizaines de kbps à 1 Mbps
- vidéo à la demande de qualité au moins égale à la VHS basée sur le MPEG1, MPEG2 et MPEG4 allant de quelques centaines de kbps à plusieurs Mbps

Plusieurs applications existent et permettent de diffuser des vidéos sur Internet, notamment Real Helix producer de RealNetworks, le WindowsMediaServer de Microsoft, Darwin d'Apple et VideoLanClient de l'Ecole Centrale de Paris. Les résultats que nous ferons figurer dans ce document sont basés sur cette dernière application. La particularité de VideoLanClient est le support d'un grand nombre de formats vidéos, des protocoles IPv6 et IPv4 ainsi que la possibilité de diffusion unicast et multicast.

Les profils de VoD que nous avons capturés à partir de l'outil tcpdump se basent sur 3 vidéos aux caractéristiques différentes. Ces dernières sont rassemblées dans le Tableau 8 et leurs profils de débit sont illustrés Figure 59. Nos observations ont montré que les débits résultants sont directement fonctions du type de média à diffuser.

Tableau 8 : Caractéristiques des vidéos capturés

Fichier	MPEG	XVID	DIVX
Vidéo			
Codec	MPEG-1	Xvid	DivX 3 Low-Motion
Largeur x Hauteur	320*240	672*288	360* 288
Taux de rafraîchissement des images (fps)	30	24	25
Type de débit	CBR	VBR	VBR
Débit (kbps)	1035	1274 (en moyenne)	297 (en moyenne)
Audio			
Codec	MPEG-1 Layer 2	MPEG-1 Layer 3	PCM Audio
Taux d'échantillonnage (Hz)	44	44	44
Débit (Kbps)	128	319	353

Les profils de trafic générés par VLC sont directement dépendants de la technique d'encodage vidéo utilisée. Ainsi une vidéo MPEG-1 génère un profil relativement constant sous VLC tandis que, pour des vidéos encodées en MPEG-4, la nature de ces profils s'apparente à ceux de la visioconférence. Toutefois, l'amplitude des variations et la variation elle-même se révèlent beaucoup plus importantes ([400kbps – 3.5MBps]) que pour une simple visioconférence si l'on considère une vidéo de haute qualité encodée à un débit variable au format XVID. La taille des paquets est, quant à elle, constante du fait du mécanisme d'encapsulation de VLC qui multiplexe les flux vidéo et audio en un flux MPEG2-TS qui est ensuite segmenté en paquet UDP de taille identique (par défaut à la MTU d'Ethernet mais configurable au niveau de l'application).

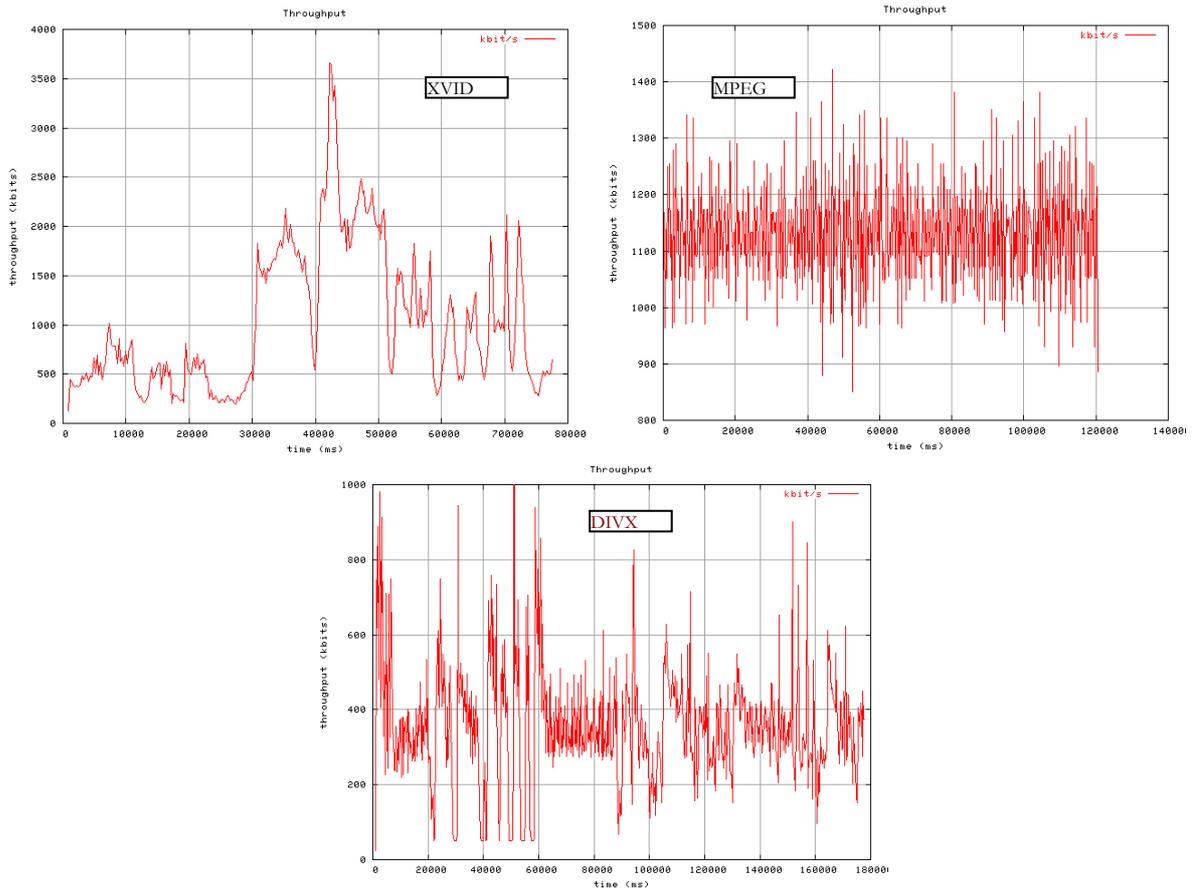


Figure 59 : Profil de débit des Sessions vidéo à la demande

Bien que l'importance des débits observés nous incite à considérer ce type de trafic comme provenant plutôt de la gateway satellite (ressource DVB-S et non DVB-RCS), il est tout de même légitime de se demander si la mise en place d'une classe de service dédiée à la diffusion de ce type de média au niveau des STs permettrait de favoriser l'apparition de diffuseurs de programmes radiophoniques ou télévisuels derrière les STs. Les progrès réalisés en terme de compression vidéo permettent de l'envisager. Cependant ce service nécessite un contrôle d'admission et une mise en place dynamique du service.

En premier lieu, la calibration du conditionnement effectué par le service associé se doit d'être configuré très précisément si l'on veut éviter une dégradation des services déjà existant avant le démarrage de la diffusion. Ce type de configuration est envisageable notamment pour tout ce qui est contenus vidéos préenregistrés dont on connaît, à l'avance, les caractéristiques.

Par ailleurs, du fait des contraintes en terme de délai et de gigue plus souples pour ces types de médias (hormis les fonctions magnétoscopes) pouvant tolérer des délais de l'ordre de plusieurs secondes, les services de type AF sont tout à fait indiqués.

Ainsi, une gestion optimisée de la qualité de service et donc de la gestion des ressources passe par un algorithme de marquage prenant en compte l'importance des différents flux et données constituant les fichiers MPEG2 et a fortiori MPEG4. L'audio apparaît généralement comme plus important que la vidéo. A l'intérieur du flux audio, la parole a également une importance plus grande que la musique. Et finalement, l'encodage hiérarchique de la vidéo permet d'offrir un traitement différencié pour les flux « principaux » (indispensables au décodeur) et les flux optionnels importants pour les couches d'amélioration vidéo qui se verront attribués respectivement un code AF « low drop precedence » et « high drop precedence » [131] [132][133].

IV.2.2.4 Les applications de données interactives

Actuellement le volume de données véhiculées par les applications multimédias reste minoritaire. Les applications de données sont elles majoritaires et se diversifient rapidement. La montée en puissance du trafic Peer-to-Peer ces dernières années le démontre.

La composition en terme de volume de trafic est bien évidemment dépendante du type de réseau considéré (académique ou commercial, ADSL ou satellite) et elle a considérablement évolué ces dernières années, le P2P ayant détrôné le trafic Web. De nombreuses études ont été effectuées pour caractériser ces différentes familles. La Tableau 9 correspond ainsi à la décomposition du trafic capturé à la sortie des plaques ADSL [134].

Tableau 9: Composition du trafic ADSL par application

	Applications	Pourcentage
Non p2p	http	14,6
	ftp	2,1
	nntp	1,9
	autres	31,8
	total trafic non p2p	50,4
p2p	Edonkey	37,5
	Kazaa&Morpheus	7,8
	Napster	3,8
	Gnutella	0,5
	total trafic p2p	49,6

Les flux issus des applications de type Telnet sont constitués de petits datagrammes (50 octets) et le serveur ne répond que de manière occasionnelle aux commandes transmises par l'utilisateur distant. Ainsi le trafic résultant est en moyenne 20 fois plus important que le trafic utilisateur issu du serveur [135]. Cependant le trafic généré par un être humain est naturellement limité par sa vitesse de frappe qui est rarement plus rapide que 5 caractères par seconde [136] ce qui donne un minimum de 200ms de délai inter-paquet. Ce trafic est donc faible en débit et génère très peu de rafales. Cependant il reste très vulnérable aux erreurs et l'on considère que 200ms constitue le seuil au-delà duquel la qualité en terme d'interactivité perçue par l'utilisateur est dégradée. Ce type d'application est de ce fait déconseillé sur les systèmes satellites géostationnaires qui ne peuvent assurer de tels délais.

Enfin le trafic Web véhiculé par HTTP sur TCP est directement lié au contenu des pages Web et à la dynamique de TCP. Des études de traces [137] ont montré que la majorité des requêtes HTTP ne dépassent pas 500 octets. Les réponses quant à elles sont typiquement de l'ordre de moins de 50 Ko mais peuvent se révéler beaucoup plus importantes si HTTP est utilisé pour télécharger des fichiers de taille conséquente à partir de pages Web. On différencie ainsi une utilisation interactive de HTTP et non interactive ou *FTP-like*. Généralement on considère que des temps de chargement de l'ordre de 5s contentent l'utilisateur. En plus d'un temps de réponse faible, il faut également assurer une faible variance du temps de réponse car l'utilisateur y est également sensible.

Bien que l'on ait différencié d'un côté Telnet et HTTP interactif et de l'autre FTP, P2P et un HTTP *FTP-like*, on peut envisager que ces deux types de trafic partagent la même file à condition d'attribuer des niveaux de rejet différents en fonction de leur niveau d'interactivité. Cependant, Telnet et les requêtes HTTP interactives devraient bénéficier de leur service propre (priorité équivalent à la VoIP) et de tailles de file réduites pour réduire les délais d'attente.

Les principales conclusions sont, dans le type de réseau satellite considéré et en particulier pour utiliser au mieux les ressources de la voie montante, qu'il faut protéger des applications de type Web interactif et Telnet du trafic FTP et P2P.

IV.2.3 Les limites

Cette caractérisation doit évidemment évoluer avec les applications. Elle ne peut que servir de base à une configuration à priori qui doit être confirmée par l'expérimentation voire l'exploitation

réelle.

Ainsi la composition du trafic interactif sur un lien satellite ne peut en toute rigueur être approchée par celle obtenue en sortie d'une plaque ADSL. En effet, les comportements des utilisateurs seront inévitablement influencés par le médium sous-jacent ne serait-ce que du fait du délai de propagation important dans les réseaux satellites géostationnaires.

Seule une analyse complémentaire de traces sur un lien bidirectionnel satellite pourrait infirmer ou confirmer une modification du profil de trafic des applications réellement interactives (Téléphonie, vidéoconférence, HTTP et SSH).

Devant la difficulté de générer un trafic fidèle aux caractéristiques des applications réelles à partir d'un modèle, la méthode que nous avons retenue afin d'évaluer les services de QoS de notre architecture est de rejouer des traces d'applications réelles. Pour ce faire, nous avons développé un outil de « rejeu » de trafic pour la plate-forme d'émulation satellite. Cet outil est ainsi à même de rejouer un trafic réseau issu de traces réelles capturées sur un véritable réseau n'importe quand et dans n'importe quel environnement compatible avec cet outil.

Nous détaillons les outils de capture et de rejeu de traces ainsi que la méthode de mesure adoptés sur la plate-forme dans la section qui suit.

IV.2.4 Outil de rejeu et plate-forme de mesures

Avant de décrire les outils de capture et de rejeu de trafic, nous allons brièvement décrire les métriques que nous avons retenues et la plate-forme de mesures de SATIP6.

IV.2.4.1 Métriques de QoS Applicative

Afin d'évaluer les performances de notre architecture de QoS, nous allons mesurer son impact sur les profils de trafic auxquels elle sera soumise en fonction des :

- conditions de charge du réseau satellite
- conditions de charge du ST
- configurations des services de QoS de la plate-forme SATIP6

Les mesures sont effectuées de l'interface réseau de la source de trafic à l'interface réseau d'entrée du récepteur. Elles visent donc à rendre compte de la QoS observée de bout en bout entre interfaces réseaux à travers l'évolution des paramètres de délai, de gigue et de taux de perte associés aux paquets d'une même connexion.

IV.2.4.2 Plate-forme de démonstration

L'objectif est de reproduire, grâce aux outils détaillés dans la section suivante, des profils de trafic multimédia et d'étudier l'impact du réseau satellite sur la transmission de ces données à partir des mesures définies précédemment. Pour ce faire, il nous faut un serveur émettant les données vers un client, les deux entités étant synchronisées temporellement à la milliseconde près par le protocole NTP déployé sur un LAN Ethernet parallèle et indépendant du réseau satellite émulé. Ainsi, après avoir évalué le comportement des applications de bout en bout par service implémenté, on peut en conclure le comportement de chaque classe de service.

La plate-forme de démonstration est illustrée Figure 60.

L'intérêt de cette plate-forme est que le segment satellite est émulé. Contrairement à la simulation qui requiert la modélisation du réseau des couches applicatives aux couches physiques, l'émulation permet l'interaction avec de véritables applications et protocoles de transport existants. Outre l'aspect démonstratif, les applications peuvent être testées en conditions réelles ce qui est un avantage incontestable pour tester les effets du réseau satellite sur leur fonctionnement et notamment leur robustesse.

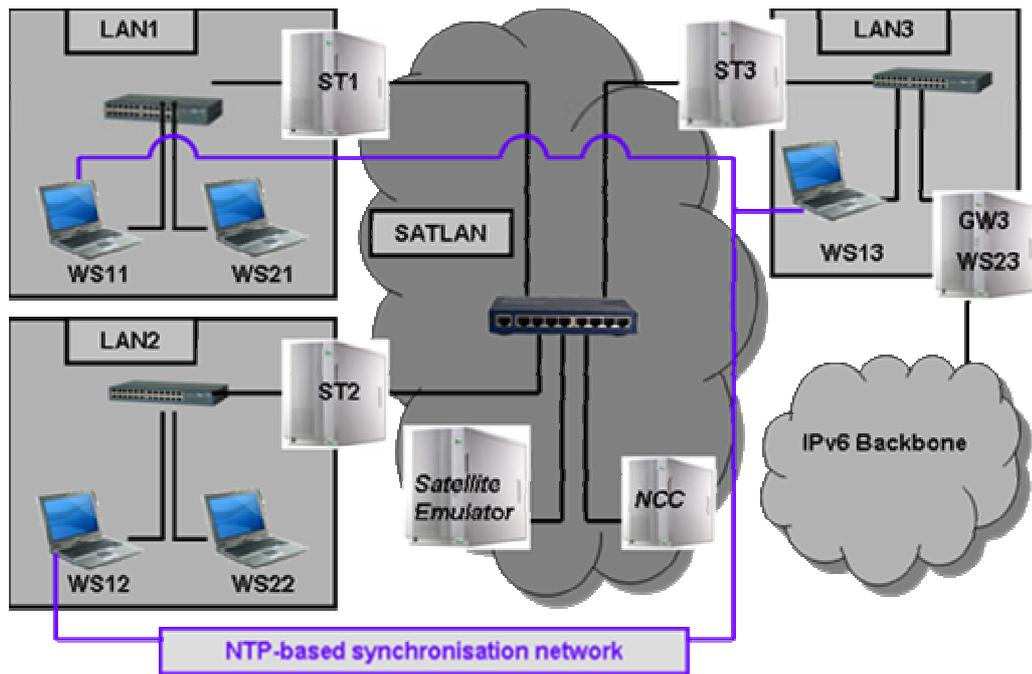


Figure 60 : Plate-forme de test

IV.2.4.3 Outils de Capture et de Rejeu

IV.2.4.3.1 Capture des flux

Partant du principe que les applications multimédias actuelles ne disposent pas de mécanismes leur permettant d'adapter dynamiquement leur débit à la charge du réseau, on peut caractériser le comportement de l'application par un nombre limité de critères.

Le premier outil, Floc (Flow Capturer) est un logiciel permettant de capturer « à la volée » pendant une durée donnée l'ensemble des données véhiculées par toutes les connexions provenant d'une application. Cet outil réutilise ainsi les fonctionnalités du traceur de connexion implémenté pour le QoS Agent.

Les données que l'on considère comme caractéristiques du comportement d'une application multimédia sont les suivantes :

- Le nombre de connexions ouvertes
- Le protocole de transport associée à chaque connexion (TCP ou UDP)
- Le temps absolu associé au paquet capturé (dont on déduit principalement le délai inter-paquet)
- La taille de la charge utile de niveau Transport associée à ce paquet

Floc fait automatiquement le lien entre le nom de l'application et les connexions qui lui sont associés et crée un filtre tcpdump constitué de quadruplets (adresse IP source, adresse IP destination, Port source, Port destination) à partir duquel la capture du trafic s'effectuera.

Le fonctionnement de cette commande étant lié aux commandes lsof et tcpdump, elle ne peut être exécutée que sur un système de type Linux. Par ailleurs, elle doit avoir lieu sur le système hébergeant l'application.

Toutefois des scripts intermédiaires nous permettent également d'interpréter directement des traces tcpdump et de constituer des traces floc de n'importe quel profil d'application.

Enfin ces traces mono-applicatives peuvent ensuite être concaténées pour former des profils de trafic multimédia agrégés.

IV.2.4.3.2 Rejeu des flux applicatifs

L'étape suivante est la régénération de flux à partir des traces capturées ou agrégées. Pour se faire, l'outil Flore (Flow Regeneration) est composé d'un client qui transmet le trafic correspondant au fichier de trace en entrée et d'un serveur qui joue le rôle de récepteur. La régénération de flux se décompose en deux étapes.

La première constitue une phase de négociation entre le serveur et le client où le client signifié au serveur le nombre de connexions qu'il va ouvrir et le type de protocole de transport associé.

La deuxième phase est le rejeu proprement dit. A l'émission de chaque paquet, le client identifie à quelle connexion le paquet appartient, associe à ce paquet une estampille temporelle ainsi qu'un numéro de séquence, le complète par des octets de bourrage à hauteur de la taille du paquet capturé et le transmet. A la réception, le serveur crée au fur et à mesure de l'arrivée des paquets un fichier de trace par connexion.

Le récepteur étant synchronisé par NTP avec l'émetteur, il est à même d'évaluer le délai d'acheminement de chaque paquet de l'interface de l'émetteur à l'interface du récepteur à travers le réseau satellite. De même, une rupture dans la séquence des paquets transmis équivaut à la perte d'un paquet dans le réseau.

IV.3 Calibration de la plate-forme

Dans cette section, nous proposons une configuration à priori de l'architecture de QoS à partir des observations réelles effectuées dans la section IV.2.2. Nous allons nous baser sur les caractéristiques des applications pour configurer les services de QoS IP. Cette configuration est réalisée également de manière cohérente avec les ressources disponibles au niveau 2.

Une configuration de base est ainsi proposée. Elle sert uniquement de référence et ne prétend pas être la configuration optimale. Cette configuration sera certainement amenée à évoluer en fonction notamment des résultats des évaluations menées en section IV.4 afin d'optimiser la gestion des ressources en fonction du type de trafic auquel elle sera soumis.

IV.3.1 Remarques préliminaires

Auparavant nous rappelons les formules de conversion entre les débits au niveau IP et ATM nécessaires pour assurer une cohérence entre les ressources attribuées par le DAMA et celles partagées au niveau IP.

IV.3.1.1 Mise en correspondance des débits ATM et débits IP

La conversion d'un débit ATM en un débit IP et inversement ne peut se faire sans connaître la taille des paquets IP. Cette conversion se révèle essentielle puisque le surcoût protocolaire est différent selon la taille de paquet IP considérée.

La Figure 61 illustre ainsi l'encapsulation d'échantillons audio dans des PDUs RTP jusqu'à l'encapsulation dans les cellules ATM.

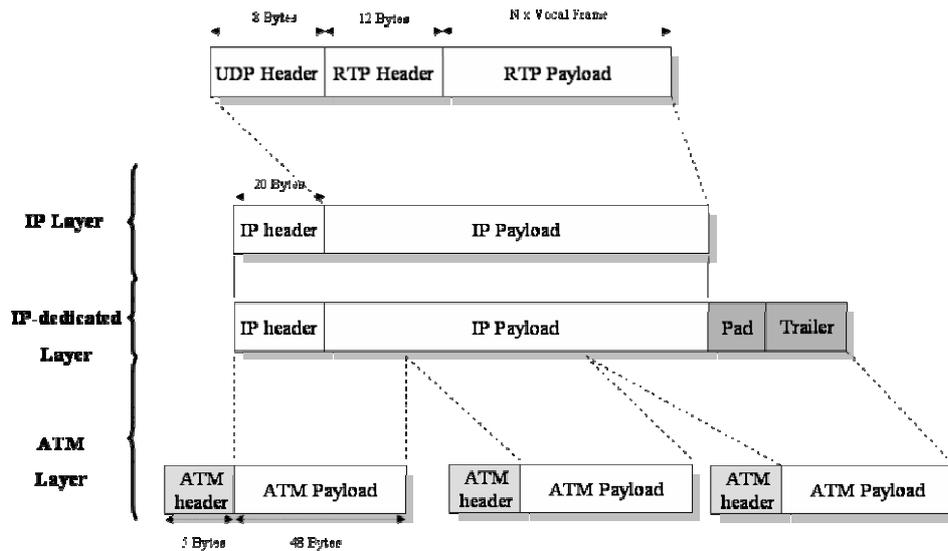


Figure 61 : Encapsulation RTP/UDP/IP/AAL5/ATM

Ainsi si l'on se réfère aux applications de Voix sur IP émettant des débits de type CBR (Constant Bit Rate), on peut les convertir en débits ATM selon les Formules 1 et inversement selon les Formules 2. Nous rappelons que la taille du Trailer AAL5 est de 8 octets

Formules 1 : Débit ATM à partir du débit du codec audio

$$PPS = \frac{\text{Débit_Codec}}{\text{Taille_Données_RTP_Utiles}}$$

$$\text{Débit_RTP} = PPS \times (\text{Taille_Données_RTP_Utiles} + 12)$$

$$\text{Débit_UDP} = PPS \times (\text{Taille_Données_RTP_Utiles} + 20) = PPS \times (\text{Taille_Données_UDP_utile} + 8)$$

$$\text{Débit_IP} = PPS \times (\text{Taille_Données_RTP_Utiles} + 40)$$

$$\text{Débit_ATM} = PPS \times \left(\left(\left\lceil \frac{\text{Taille_Paquet_IP} + \text{AAL5_Trailer}}{48} \right\rceil + 1 \right) \times 53 \right) = \text{Débit_IP} \times \left(\frac{\left(\left\lceil \frac{\text{Taille_Paquet_IP}}{48} \right\rceil + 1 \right) \times 53}{\text{Taille_Données_RTP_Utiles} + 40} \right)$$

Formules 2 : Débit IP et RTP utile à partir d'un débit ATM

$$\text{Nombres_Cellules_ATM_correspondantes} = \left(\left\lceil \frac{\text{Taille_Paquet_IP} + \text{AAL5_Trailer}}{48} \right\rceil + 1 \right)$$

$$PPS = \frac{\text{Débit_ATM}}{53 \times \text{Nombres_Cellules_ATM_correspondantes}}$$

$$\text{Débit_IP} = PPS \times (\text{Taille_Donnée_UDP_utile} + 28)$$

$$\text{Débit_utile_UDP} = PPS \times (\text{Taille_Donnée_UDP_utile})$$

IV.3.2 La configuration de référence

IV.3.2.1 Couche Physique

La configuration de la couche physique est équivalente à celle suggérée au chapitre 3 à la section 4.4 et est résumée dans le Tableau 10.

Tableau 10 : Configuration de référence de la couche physique

Couche Physique DVB-RCS						Emulateur Satellite		
Débit crête de transmission ST	Supertrame	Durée Trame	Ressource Globale	Période des CRs	Période du TBTP	Délai	Gigue	BER
[144kbps, 384kbps, 1024kbps, 2048kbps]	10 Trames	50ms	Fonction du scénario	500ms	500ms	250ms	$\pm 1ms$	Conditions météorologiques favorables

IV.3.2.2 Couche MAC

Les principaux paramètres sont liés au DAMA et au partage des ressources au sein du ST entre les deux classes de services MAC (cf. Tableau 11). Le débit de transmission globale du ST est partagé entre ces deux classes. Le CRA est d'abord défini, ensuite le reste des ressources est déterminé à travers le facteur d'anticipation α et le FCA défini comme un seuil de débit maximum alloué par ST. Par défaut, α est à 1 (aucune anticipation considérée) et le NCC ne pourvoit aucune autre capacité que celles requises (FCA = 0 kbps). Enfin le VBDC dépend directement de la configuration de l'algorithme de DAMA.

Tableau 11 : Configuration de la couche MAC

NCC		SLA ST	
FCA	α	CRA	VBDC
0 kbps	1	Multiple de 16kbps	Dépend directement de l'algorithme de DAMA

IV.3.2.3 Couche IP

Au niveau IP, la plate-forme donne accès à la configuration du nombre de classes de service supportées par le ST. Les paramètres de l'ordonnanceur sont également accessibles. Enfin un conditionnement de trafic basé sur l'algorithme du Token Bucket est associé à chaque classe de service, excepté le service Best-Effort. Ce contrôle d'admission statique permet de répartir la bande passante disponible entre les différentes classes de service conformément au SLA du ST.

IV.3.2.3.1 Token Bucket

Le conditionnement du trafic pour les services EF et AF a lieu à travers un Token Bucket. Il se présente de la manière suivante :

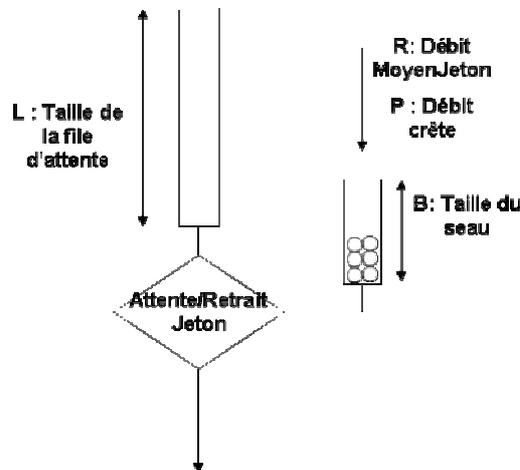


Figure 62 : Algorithme Token Bucket

Les paramètres sont :

- R, le débit moyen : il permet de fixer le débit moyen à long terme d'émission des paquets.
- P, le débit crête : il permet de fixer le débit maximum auquel peut être envoyé un paquet dans un intervalle de temps court
- B, la taille du seau : il constitue la taille maximale de la rafale de paquet qui peut être envoyé
- L, la taille de la file d'attente : elle constitue le nombre d'octets maximum qui peut être mis en file d'attente en attendant la disponibilité des jetons

IV.3.2.3.2 Configuration d'un service de Voix sur IP de type EF

Dans le chapitre 3, nous avons précisé que le conditionnement EF se basait sur un Token Bucket avec B fixé à la taille maximale des paquets de l'application et R correspondant au CRA converti au niveau IP (si l'on ne considère ce service de Voix sur IP comme l'unique service EF).

Supposons maintenant que ce service soit à même de supporter quatre connexions GSMs simultanées. Une connexion GSM initiée par Gnomemeeting est caractérisée par des paquets RTP de charge utile égale à 132 octets et le débit du codeur est de 13.2 kbps. Le débit IP correspondant est donc de $\frac{13200}{132} \times (132 + 40)$ soit 17.2 kbps. Le paquet IP auquel on ajoute les 8 octets du trailer AAL5 (soit 180 octets) est fragmenté en 4 cellules ATM. Le débit ATM équivalent est donc de 21.2 kbps. Le débit total au niveau ATM est de 84.8 kbps. La granularité d'attribution des ressources au niveau MAC étant de 16 kbps, le CRA associé est de 96 kbps. Au niveau IP, ce débit équivaut à 77.9 kbps

La configuration du Traffic Conditioning Block (TCB) associé, si l'on considère qu'il ne conditionne que des paquets provenant de codeurs GSM, est donc la suivante :

- La taille du seau B est de la taille maximale du paquet envoyé par l'application, soit 172 octets.
- Le débit des jetons est de 77.9 kbps

IV.3.2.3.3 Configuration du niveau IP

Dans notre cas, les TCBs du niveau IP sont configurés selon les paramètres figurant dans le Tableau 12.

Tableau 12 : Configuration de la QoS de la couche IP

Service	EF	AF	BE
Taille du seau	172 octets (1 GSM packet size)	1500 octets (Ethernet MTU)	No Conditionning
Débit moyen	77.9kbps (corresponds to a 96kbps CRA)	Fonction du débit de transmission du ST	
Taille File d'attente	688 octets (4 paquets GSM)	Fonction des applications	500000 bytes
Délai EDF			
Latence maximum	25ms	500ms	5s

Le « délai EDF » constitue l'échéance au bout de laquelle le paquet doit être envoyé et la « latence maximum » est le temps maximum que le paquet peut attendre sans être servi par l'ordonnanceur avant d'être rejeté.

Les services EF et AF peuvent être décomposés en plusieurs services pour différentes catégories de trafic. Nous noterons notamment la nécessité de fournir un service *EF-sig* pour la signalisation, et un autre service *EF-vidéo* pour la vidéo dans des services de type vidéoconférence. De même, le service AF peut être décomposé en plusieurs services. On peut ainsi envisager un service dédié au streaming vidéo, un autre aux applications de données interactives telles que

HTTP...

Un service EF et un service AF sont cependant suffisants pour démontrer l'intérêt de l'architecture de QoS proposée.

IV.4 Evaluation de QoS

Plusieurs tests ont été menés afin de valider l'architecture de QoS et leurs résultats sont répertoriés dans cette section.

La plupart des résultats seront expliqués à partir de fonctions de distribution cumulative des délais ou du délai et de la gigue en fonction du temps.

En ce qui concerne le délai et la gigue, la précision obtenue par l'intermédiaire du réseau parallèle NTP synchronisé par GPS est de l'ordre de la ms. La reproductibilité des expériences est assurée par le rejeu de traces de trafic identiques d'un scénario à un autre afin de montrer l'influence de chaque paramètre de configuration de la plate-forme sur les performances des applications.

Enfin la congestion au sein d'un ST est émulé par l'intermédiaire d'un générateur de trafic Iperf [138] qui génère un trafic parasite.

IV.4.1 Comportement de TCP

Le premier test vise à évaluer l'influence du DAMA sur le comportement de TCP. Le test consiste au transfert par FTP d'un fichier de 1 Gbit à travers le service BE de la station WS11 à la station WS12 (cf. Figure 60). TCP a été préalablement configuré en fonction de la forte latence induite par le réseau satellite (cf. Tableau 13).

Tableau 13 : Adaptation de la pile TCP au segment satellite

Maximum Segment Size (bytes)	1460
Receive Buffer (bytes)	128000
Delayed ACK Mechanism	Segment/Clock Based
Maximum ACK Delay (sec)	0.2
Slow-Start Initial Count (MSS)	2
Fast Retransmit	Enabled
Fast Recovery	Enabled
Window Scaling	Enabled
Selective ACK (SACK)	Enabled
Nagle's SWS Avoidance	Disabled
Karn's Algorithm	Enabled
Retransmission Thresholds	Attempt Based
Initial RTO (sec)	2
Minimum RTO (sec)	1
Maximum RTO (sec)	30
RTT Gain	0.125
Deviation Gain	0.25
RTT Deviation Coefficient	4

La Figure 63 montre l'évolution du débit TCP et des CRs associées. Dans ce scénario, le FCA a été placé à 80 kbps par ST et $\alpha = 0$. Les ressources globales disponibles pour les classes DVB-

NRT est de 1.5 Mbps.

Avant de décrire cette expérience, nous noterons que le FCA joue ici un rôle primordial dans l'acheminement des acquittements (ACK) liés à la connexion TCP. Les ACKs, qui constituent le feedback nécessaire à TCP pour augmenter la taille de sa fenêtre de congestion, sont transmis par la station réceptrice WS12 à travers le service Best-Effort. Dans notre cas de figure, si la station WS12 est la seule à émettre sur le LAN2 vers le réseau satellite, des tests ont montré que sans FCA les performances de TCP sont réduites. Ce phénomène s'explique par la sporadicité du trafic des ACKs qui ne suffit pas à lui seul à engendrer des requêtes de capacité régulières auprès du NCC par le ST2. Les transmissions de certains ACK sont alors retardées par la latence moyenne induite par le DAMA soit 1250ms (cf. section III.3.3.4.2) et le débit de la connexion TCP en est directement affecté. La capacité allouée à travers le FCA permet de se prémunir contre ce type de phénomène lorsqu'un ST est particulièrement sous chargé. Par ailleurs, si le trafic des ACKs est multiplexé avec d'autres trafics, il profite alors des CRs des autres flots ce qui permet de réduire leur temps de séjour dans la queue MAC à condition que le ST n'entre pas en congestion.

Nous pouvons maintenant commenter l'expérience dans son ensemble. Le scénario adopté est le suivant. A $t = 0s$, 3 STs sont connectés, ST_1 qui requiert 100 kbps en continu (flux UDP) en plus du flux TCP, ST_2 qui requiert 100 kbps en continu (Flux UDP + ACK) et ST_4 (un ST virtuel¹) qui requiert 800 kbps en continu. A $t_1 = 60ms$, ST_5 se logge et requiert 400 kbps. Enfin à $t_2 = 120ms$, ST_6 se logge et requiert 200 kbps. A $t_3 = 420ms$, ST_4 , ST_5 et ST_6 se délogent.

Entre t_2 et t_3 , se produit une congestion sur le lien montant dans le réseau satellite puisque chaque ST s'il transmettait au maximum de ses capacités devrait recevoir 300 kbps chacun (soit 1.5 Mbps/5 stations). Cependant ST_2 et ST_6 à eux deux ne requièrent que 300 kbps. Il reste donc 1.2 Mbps à partager entre les trois autres STs, soit 400 kbps chacun. La Figure 63 montre que malgré cela la connexion TCP arrive à accaparer 350 kbps.

Quand la congestion se résorbe, c'est-à-dire quand les STs se délogent à t_3 , on remarque que que le ST_1 profite de la bande passante nouvellement disponible. En exploitant le FCA, le DAMA permet aux paquets de la connexion TCP d'atteindre de faible temps de séjour dans la queue MAC en affichant un RTT de moins de 0.6s. TCP est alors en mesure de transmettre à approximativement 900 kbps. On peut remarquer que les CRs correspondent parfaitement au débit TCP. En effet, la charge utile d'une cellule ATM étant de 384 bits, et la durée de la supertrame de 0.5s, 1.3 cellule ATM par trame équivaut à 1 kbps.

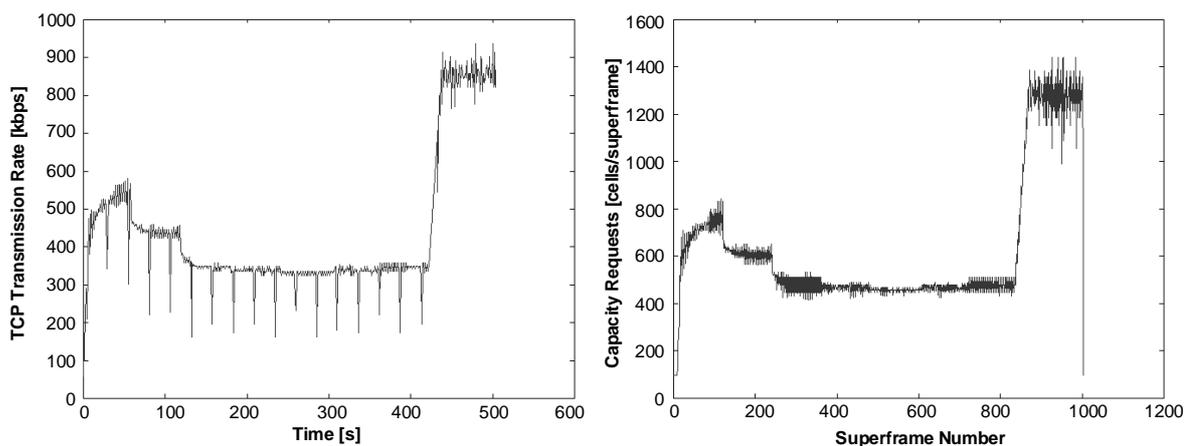


Figure 63 : (a) débit TCP durant la congestion, (b) requêtes de capacité associées

¹ Au sein du NCC, il est possible de jouer des requêtes de capacité virtuelles associées à un ST virtuel à partir de traces de CRs préétablies ou enregistrées au cours de sessions réelles. Par défaut, tous les STs avec un identifiant supérieur à 3 (nombre de STs réels implémentés dans la plate-forme), $ST_{x>3}$, sont considérés comme virtuels.

IV.4.2 L'influence du DAMA sur les sources à débit constant et variable : configuration du paramètre d'anticipation

Dans les expériences de cette section, nous ne nous plaçons pas dans des situations de congestion ni du ST ni du lien montant afin d'évaluer l'influence seule du DAMA.

Nous voulons tout d'abord démontrer que le DAMA ne se comporte pas de la même manière vis-à-vis d'un flux CBR et d'un flux VBR (Variable Bit Rate). Même si cela paraît étrange au premier abord, un flux VBR se comporte mieux qu'un flux CBR si le DAMA est dans sa configuration de base. Nous allons expliquer ce phénomène et proposer une configuration du DAMA adapté au flux CBR.

Les premières observations sont les suivantes. A $t = 0s$, un flux CBR démarre : un flux GSM (sans suppression de silence) dans la Figure 64 et un flux MPEG (cf. Tableau 8) dans la Figure 65.

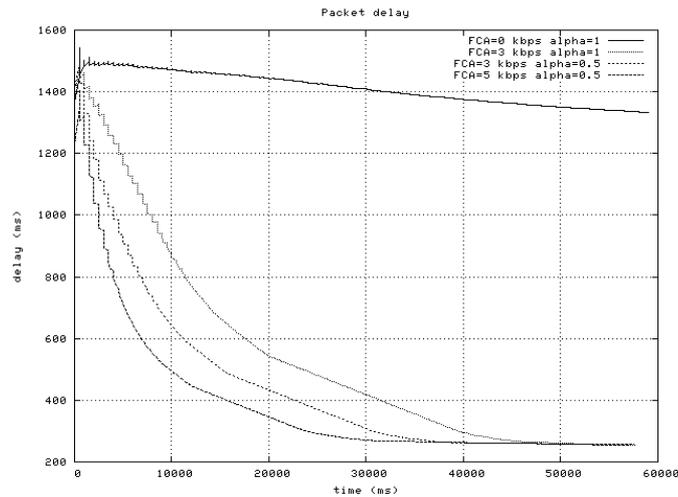


Figure 64 : Influence du DAMA sur une connexion GSM

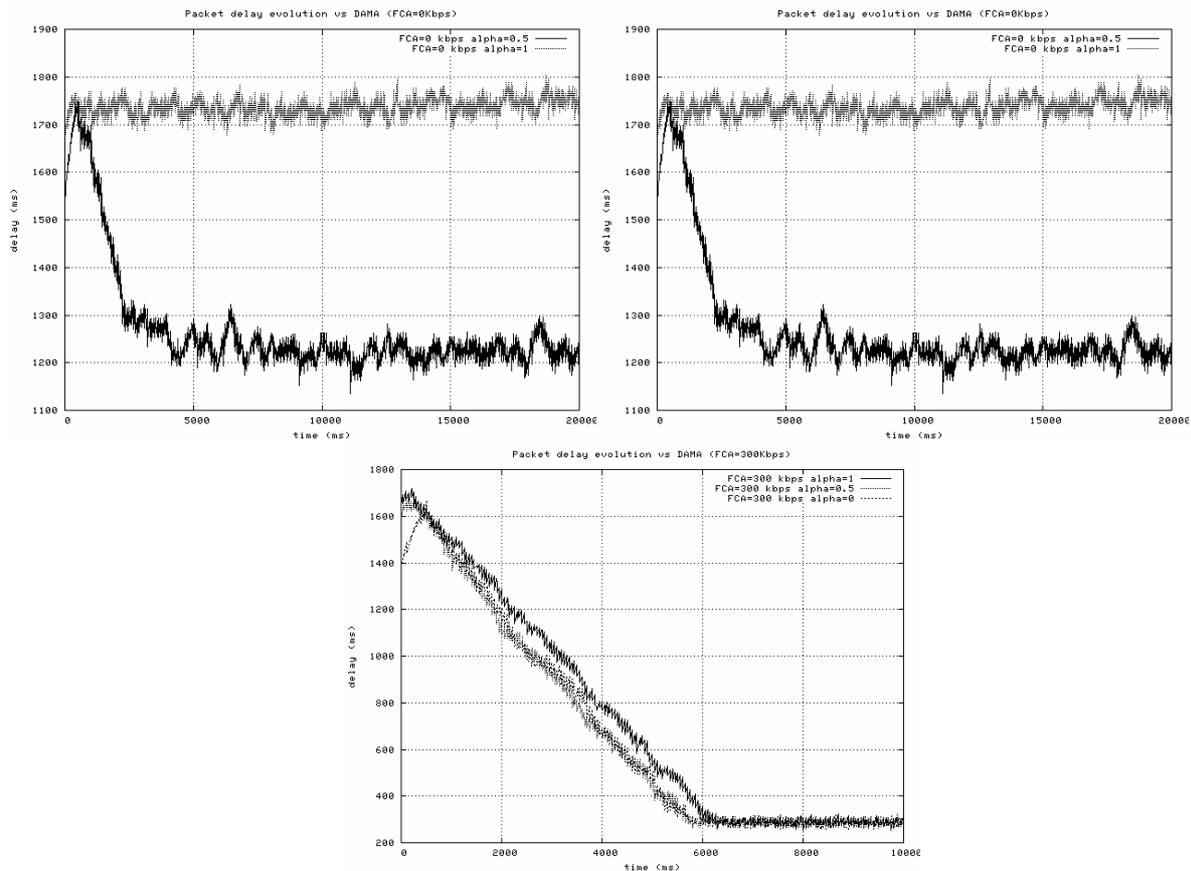


Figure 65 : influence du DAMA sur un profil MPEG (a, b &c)

Le DAMA dans sa configuration de base n'est pas à même d'absorber convenablement le pic de débit d'un flux CBR :

- Pour un FCA=0 et quelle que soit la valeur de α , le délai des paquets est toujours supérieur à 1250ms.
- Pour des FCAs supérieurs, la taille de la file MAC DVB-NRT est progressivement réduite à zéro grâce à la capacité allouée par le FCA. Le processus est cependant relativement long.

Ce phénomène s'explique par le fait que les premiers paquets sont accumulés à un débit constant durant le délai introduit par le DAMA. Ensuite les requêtes VBDC atteignent rapidement un régime stationnaire calqué sur le débit CBR de l'application qui n'est pas suffisant pour éliminer le surplus emmagasiné en début de session. Le paramètre d'anticipation α couplé avec le FCA (Figure 65 c) permet clairement d'atténuer ce phénomène.

L'étude suivante porte sur l'influence du DAMA sur une source de trafic VBR : les caractéristiques de la source sont données dans le Tableau 8 sous le profil DIVX (Figure 66 a)). Les Figure 67 a) et b) montrent les fonctions de densité cumulative des délais par paquet de bout en bout pour des FCA de 0 et 40KBps (10% du débit moyen de la vidéo) et pour α entre 0 et 1. Ces figures démontrent que l'architecture proposée avec un $\alpha = 0.5$ améliore significativement les performances de délai comparativement à l'architecture de référence où $\alpha=1$. De plus, il n'y a aucun avantage à réduire α . Le délai moyen au cours des expérimentations est de 1s pour α égal à 0.5 ou 0 et de 1.7s pour $\alpha = 1$. Si le FCA est conjointement pris en compte, le délai est également réduit mais l'influence du α semble atténuer. Toutefois le ST et la liaison montante n'étant pas congestionnés, le FCA ici est assimilable à du CRA puisque les capacités restantes sont distribuées entre les STs concurrents à chaque supertrame sans qu'ils aient à en faire la demande.

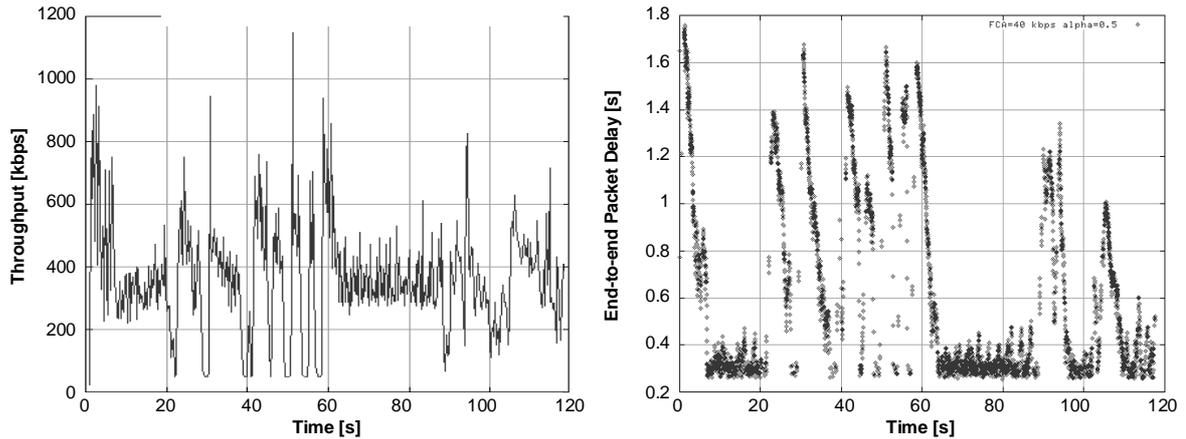


Figure 66 : a) Débit d'une vidéo à la demande basé sur le codec DIVX (Débit moyen = 400 kbps) b) délai de bout en bout par paquet avec $\alpha=0.5$ et FCA=40 kbps

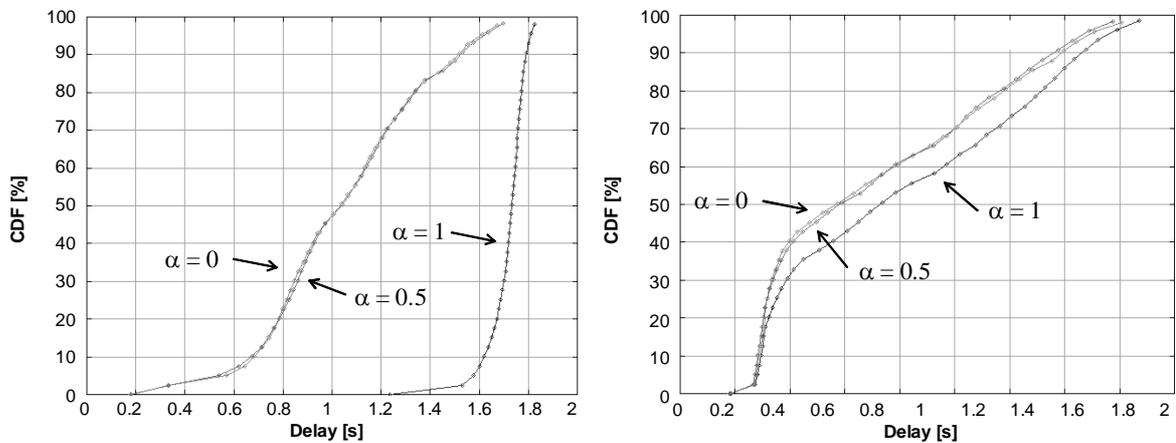


Figure 67 : a) Fonction de densité cumulative des délais de bout en bout avec un FCA = 0 kbps b) Avec un FCA = 40 kbps

A travers ces différents exemples, nous avons mis en évidence que l'efficacité du DAMA est fortement liée à la vitesse et à l'amplitude des variations du débit du ST. Ainsi, les profils de trafic des applications VoIP dotés d'un détecteur de silence sont assimilables à des processus On-Off. La réactivité du DAMA n'est pas suffisante pour pouvoir absorber des variations aussi nettes de débit. Par contre, des débits oscillants rapides et de faible amplitude ou des variations de débit de forte amplitude mais étalées dans le temps correspondent aux types de trafic pour lesquels le DAMA est tout à fait pertinent. Il adapte ainsi dynamiquement le partage des ressources satellites en fonction des besoins de chaque ST tout en offrant une QoS répondant aux contraintes de ces applications. La réactivité du DAMA peut être accrue par l'intermédiaire du facteur d'anticipation α et il peut être efficacement complété par un mécanisme tel que le FCA.

IV.4.3 Evaluation de la QoS offerte par un service de type EF

Les applications de Voix sur IP sont les plus exigeantes en terme d'interactivité. Par ailleurs, elles sont généralement configurées par défaut pour détecter les silences. Leur comportement peut alors être approché par une source de trafic On-Off, en maintenant un débit CBR durant les périodes d'activité du correspondant et un débit quasi-nul durant ces silences. Comme nous l'avons vu précédemment, l'allocation de ressource par le DAMA n'est pas assez réactive pour répondre à de brusques variations de débit. Il est alors naturel de l'orienter vers le service EF qui bénéficie d'une allocation de ressource plus stricte à travers le CRA.

Le service EF SATIP6 dans sa configuration de base est configuré pour accepter 4 connexions GSM simultanées. Par conséquent, dans un premier temps, nous avons évalué l'impact des codecs audio sur les délais de transmission et la gigue expérimentée dans des scénarios mono-flux où des sessions VoIP préenregistrées sont rejouées à travers le satellite sous des conditions de charge des services AF et BE différentes. Dans ces scénarios, quelque soit les codecs utilisés et la charge considérée, la QoS offerte par le service EF est préservée et elle garantit pour le flux un IPTD et IPDV strictement borné : $\text{IPTD} \in [250\text{ms}; 320\text{ms}]$ et $\text{IPDV} \in [-40\text{ms}; 40\text{ms}]$.

La suite des expérimentations a porté sur la répercussion de l'agrégation de flux homogènes au sein du service EF. Afin de vérifier que la superposition de plusieurs connexions au sein d'un même service ne nuit pas à la QoS obtenue, le service EF a été soumis à un agrégat de 4 flux GSM simultanés. D'après les Figure 68 et Figure 69, on remarque que les différents flots bénéficient statistiquement de la même QoS au sein du service EF, une QoS plus acceptable sur un lien satellite et cela indépendamment de la charge de trafic dans les services AF et BE, tant que le débit seuil du régulateur n'est pas dépassé (Contrôle d'admission).

Nous remarquerons que la répartition des délais est quasi linéaire. L'intervalle des délais et des giges observés correspond à la durée de la trame.

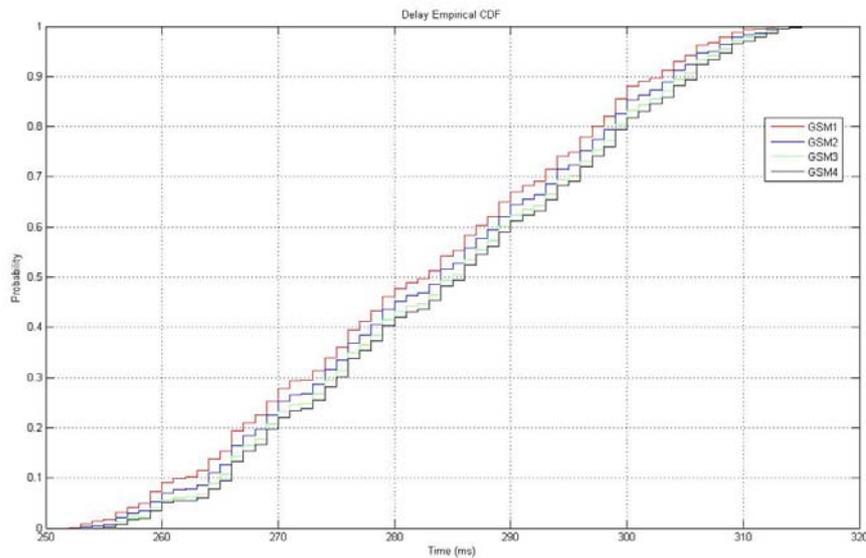


Figure 68 : Influence de l'agrégation de trafic VoIP homogène sur IPTD

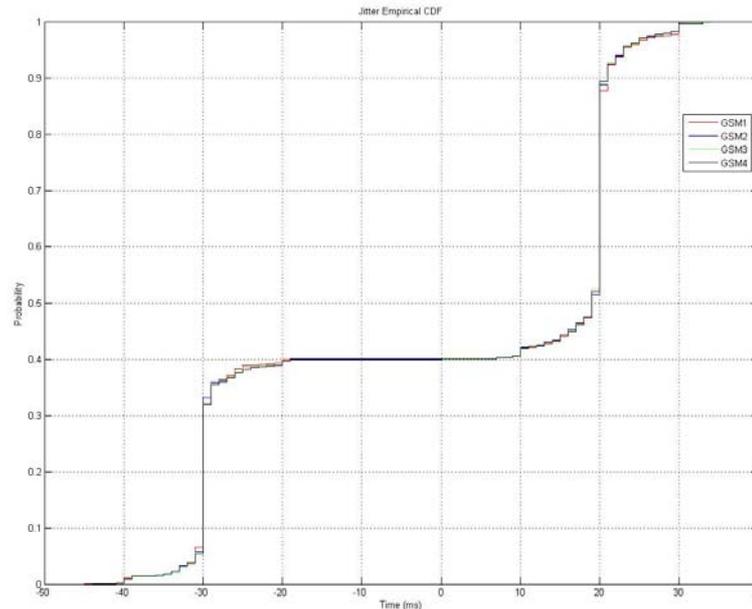


Figure 69 : Influence de l'agrégation de trafic VoIP homogène sur IPDV

IV.4.4 Evaluation globale de l'architecture de QoS

Afin de révéler au mieux l'influence de l'architecture de QoS, le réseau d'un ST va être soumis à différents niveaux de charge et à un trafic hétérogène composé de connexions GSM basées sur le codec GSM 06.10, d'une vidéo à la demande (profil DIVX) de basse qualité et du trafic FTP. La voix sur IP est à destination du service EF. La vidéo à la demande emprunte le service AF. Enfin le trafic FTP est orienté vers le service Best-Effort.

La congestion du ST est obtenue à partir d'un trafic UDP parasite généré en plus des différents trafics précédents. Ainsi, pour un taux de charge de 100%, la somme des trafics est égale au débit maximum d'émission du ST. Les mesures sont effectuées quand un régime stationnaire est établi, ce qui explique que les trafics AF et BE bénéficient d'une QoS équivalente à celle d'EF pour des taux de charge du ST inférieur à 100%. Au-delà, notre architecture revêt toute son importance puisqu'elle protège les classes de plus haute priorité de la congestion. Les différences entre les QoS offertes par chaque service selon la charge du ST sont résumées dans le Tableau 14.

Le service EF n'est pas du tout affecté par l'état de charge du ST et offre des garanties de QoS strictes en terme de délai, gigue et taux de perte.

Par contre, les services BE et AF voient leur délai moyen augmenter avec la charge du ST. Cette augmentation est toutefois maîtrisée pour le service AF puisqu'elle est conditionnée par la taille de la queue MAC DVB-NRT partagée par le trafic AF et BE soit un temps de séjour moyen des cellules issues du service AF de 1.5s.

Le service BE ne bénéficie pas d'un contrôle d'admission et n'est pas traité prioritairement par l'ordonnanceur IP. Le trafic BE s'accumule alors à hauteur de la taille de la queue BE et le surplus de trafic est rejeté d'où un taux de perte croissant en fonction de la charge.

On peut remarquer également que la gigue est limitée pour les services AF et BE : la raison réside dans le fait que les files étant toujours pleines, l'effet de « rafale » est atténué.

Ces résultats démontrent que l'architecture proposée permet un traitement différencié adapté aux contraintes de chaque classe de trafic tout en préservant par l'intermédiaire du DAMA les ressources au niveau satellite.

Tableau 14 : Performance de l'architecture de QoS SATIP6 dans différentes conditions de charge

Network Load [%]	Average Delay (Jitter) [ms]			Losses [%]		
	BE	AF	EF	BE	AF	EF
25	293 (26)	293 (26)	283 (23)	0	0	0
50	291 (23)	291 (24)	283 (23)	0	0	0
75	290 (24)	289 (24)	283 (23)	0	0	0
125	6753 (23)	1783 (24)	283 (23)	33	0	0
150	6755 (28)	1783 (24)	283 (23)	37	0	0

IV.4.5 Evaluation des performances de l'architecture de QoS basée sur de l'agent de QoS et le proxy SIP

IV.4.5.1 Agent de QoS

Le scénario suivant a été mis en œuvre afin de mettre en évidence l'influence du changement de service initié par l'agent de QoS.

Un flux vidéo UDP de 1 Mbit/s est émis en best-effort; 25 secondes plus tard, un flux perturbateur venant saturer le lien montant est mis en concurrence ; le débit du flux initial tombe à 800 kbit/s. Après 25 secondes, le service 'Voice' est appliqué au flux initial. Après un pic de débit qui correspond à la superposition du trafic issu du service EF à 1 Mbps et du trafic emmagasiné dans le service BE avant la réservation de ressource, le débit se stabilise à nouveau autour de 1 Mbits/s. A $t = 80s$, le service est repassé en best effort.

L'évolution du débit de l'application bénéficiant de ce changement dynamique de QoS est illustrée dans la Figure 70.

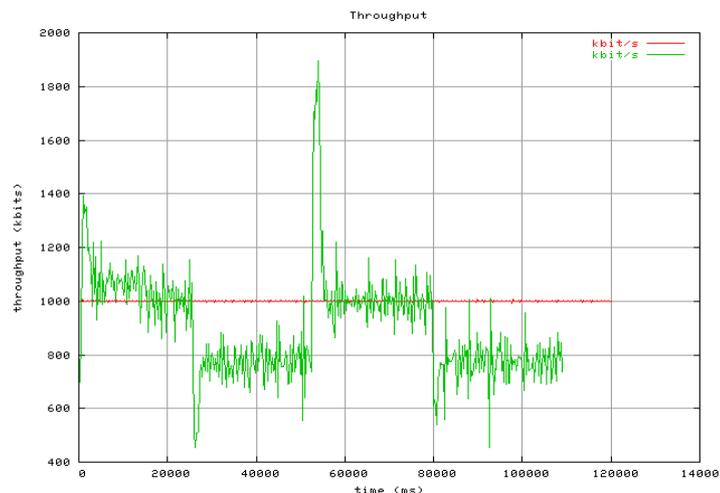


Figure 70 : Changement dynamique de Cds initié par l'agent de QoS

IV.4.5.2 Proxy SIP

Cette architecture détaillée dans le chapitre 3 permet aux différents flux constituant une session multimédia de bénéficier de la QoS associée définie par l'administrateur du réseau. Dans le cas d'une session VoIP, le flux audio bénéficiera donc de la QoS offerte par le service EF.

Dans cette section, nous nous intéresserons à l'aspect signalisation, évaluerons le temps de mise en place d'une session SIP et démontrerons l'intérêt d'une classe de service dédiée à la signalisation.

Ainsi les différents intervalles de temps s'écoulant entre les messages INVITE-RINGING, OK-ACK et BYE-OK nous permettent d'évaluer le temps d'établissement d'une session SIP. Ces intervalles ont été choisis dans la mesure où aucune interaction avec l'être humain n'est

impliqué ce qui ajouterait un délai non maîtrisé dans nos mesures. Ainsi le message OK n'est envoyé qu'au moment où le correspondant accepte la conversation (par exemple en décrochant son téléphone SIP). La Figure 71 nous permet de visualiser les différents temps caractéristiques de l'établissement d'une session SIP : $\Delta t_{\text{INVITE-RINGING}}$, $\Delta t_{\text{OK-ACK}}$, $\Delta t_{\text{BYE-OK}}$.

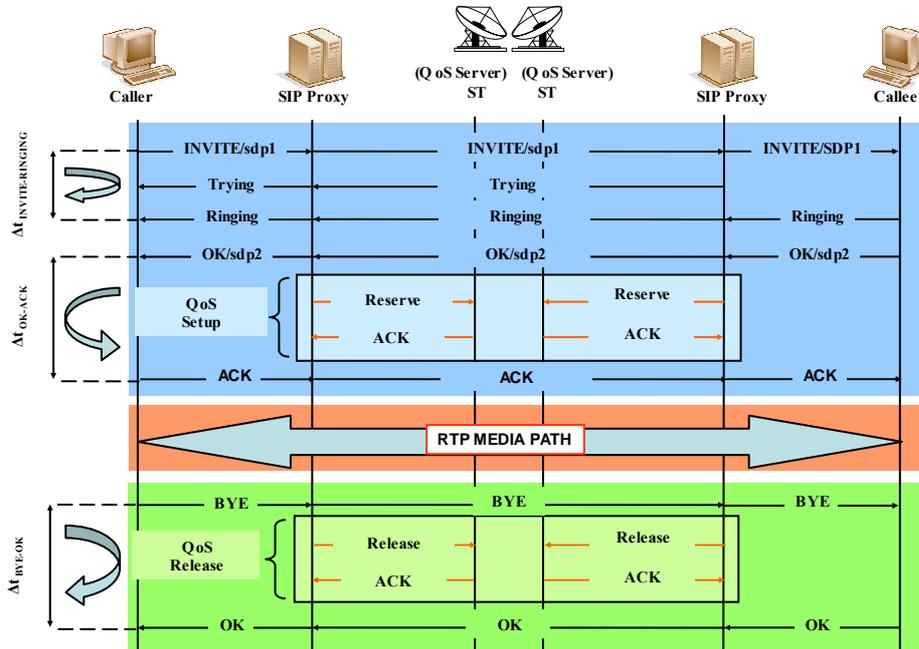


Figure 71 : Temps caractéristiques d'établissement d'une session SIP

Le Tableau 15 nous démontre que tant que le ST n'est pas surchargé et que le trafic SIP est multiplexé avec un trafic minimum les temps d'établissement de session sont relativement courts : de l'ordre de 1.5s. Ici le trafic SIP est directement véhiculé de l'appelant vers l'appelé à travers leur proxy respectif sans redirections SIP intermédiaires. Ayant lieu localement du fait de la nature distribuée de l'architecture, le contrôle d'admission et la réservation de ressource sont réalisés en quelques dizaines de millisecondes. Cependant l'impact de la congestion sur $\Delta t_{\text{INVITE-RINGING}}$, $\Delta t_{\text{OK-ACK}}$ et $\Delta t_{\text{BYE-OK}}$ n'est pas anodin puisqu'elle porte ce temps d'établissement à plus de 7s.

Le service EF est, dans ce cas de figure, beaucoup plus adapté pour le transfert de la signalisation SIP relativement réduite derrière un ST.

Tableau 15 : temps moyen d'aller-retour des messages SIP en fonction de la charge de trafic du ST

Service/ Conditions de charge	INVITE- RINGING (sec)	OK-ACK (sec)	BYE-OK (sec)
BE/60%	0.75	0.71	0.68
BE/80%	0.75	0.73	0.69
BE/100%	3.25	3.40	3.12
BE/120%	3.85	3.53	3.46
EF	0.70	0.73	0.75

IV.5 Conclusion

Au cours de ce chapitre, nous avons adopté une démarche transversale allant des couches applicatives jusqu'à la couche MAC afin d'offrir une configuration des différents mécanismes de la plate-forme qui soit cohérente avec les catégories d'applications multimédias des futurs NGNs.

Afin de reproduire fidèlement le comportement d'applications multimédias sur notre plateforme d'émulation, nous avons implémenté un outil jouant des profils de trafic multimédia typiques de plusieurs catégories d'application. De ces profils, nous avons également déduit une configuration de l'architecture de QoS.

Pour évaluer les performances de l'architecture, nous nous sommes basés sur les métriques de bout en bout de délai, de gigue et de taux de perte. Nous avons ainsi pu démontrer que le service EF fournissait statistiquement des garanties strictes de QoS équivalentes à l'ensemble des connexions acceptées.

D'autre part, nous avons mis en évidence l'importance de deux paramètres sur l'efficacité du DAMA. Cette dernière dépend directement de la variabilité et de l'amplitude des débits des profils de trafic qui lui sont soumis. Dans la mesure où ces paramètres évoluent dans des échelles de temps comparables à la latence du DAMA, l'algorithme est efficace et donc optimise la gestion des ressources tout en offrant une QoS acceptable. Pour des variations de débit trop brusques ou trop amples, la QoS offerte par le DAMA est par contre fortement dégradée. La réactivité du DAMA peut toutefois être améliorée par l'intermédiaire du paramètre α qui permet d'effectuer des requêtes de capacité issues d'une anticipation de l'évolution du trafic et cela plusieurs supertrames à l'avance.

Enfin, nous avons vérifié qu'en cas de congestion du ST, notre architecture de QoS transmettait en priorité le trafic des classes de services AF et EF au détriment du trafic Best-Effort. Nous avons également évalué les performances de l'architecture de signalisation de QoS.

Conclusion Générale

Bilan

La problématique de nos travaux concerne l'intégration des réseaux satellites dans les réseaux de nouvelle génération. Ainsi, au cours du premier chapitre, nous avons montré que l'évolution actuelle des réseaux de communication tend vers l'adoption de la technologie IP et le support de la QoS de bout en bout. Dans ce cadre, nous avons démontré que les réseaux satellites ne disposaient pas, en l'état actuel, d'une architecture de QoS suffisamment mature pour supporter l'interactivité et les besoins en terme de bande passante requis par les services multimédias large bande.

Pourtant, de nombreux mécanismes satellites sont actuellement disponibles et permettraient de faire face à ces nouveaux besoins. L'analyse des normes DVB-S/RCS a révélé que les mécanismes de gestion des ressources de niveau 2 étaient encore faiblement déployés dans les systèmes satellites opérationnels actuels du fait de spécifications trop ouvertes. De même, la sous-exploitation des technologies multifaisceaux et de l'intelligence embarquée contribue à maintenir les réseaux satellites dans leur statut de technologie de niche. L'absence d'une architecture générique capable d'unifier ces différents mécanismes compromet fortement l'essor des réseaux satellites dans l'infrastructure NGN alors qu'ils disposent de tous les atouts pour offrir une qualité de service compétitive avec les réseaux terrestres.

Dans ce contexte, notre contribution a consisté, dans un premier temps, en une spécification détaillée des méthodes d'accès aux ressources satellites de niveau 2 et, dans un deuxième temps, en leur incorporation au sein d'une architecture unifiée. Les atouts majeurs de cette architecture est qu'elle prend en compte les modèles de QoS NGN développé au chapitre 1 et qu'elle assure une coordination étroite entre les mécanismes de QoS déployés à différents niveaux de communication. Ces mécanismes comprennent notamment :

- Au niveau MAC, un contrôle d'admission orienté connexion (CAC) ainsi qu'une allocation dynamique de bande passante à la demande (DAMA)
- Au niveau IP, une architecture de type DiffServ
- Au niveau Application, des solutions de provisionnements automatiques de ressources pilotés par un agent de QoS ou couplés à l'établissement d'une session SIP.

Ce traitement hiérarchique et différencié du trafic applicatif à différents étages et selon différentes granularités a permis de partager au mieux les ressources satellites. Ainsi les applications contraintes temporellement en s'appuyant sur une signalisation applicative pilotant le provisionnement de ressources dans le réseau satellite sont à même de caractériser leurs besoins, de déclencher la mise en place des ressources correspondantes et de ne les mobiliser que durant la durée d'activité de la session. Pour les applications non contraintes temporellement, les ressources du satellite sont réparties entre les STs grâce au DAMA au niveau MAC qui alloue dynamiquement la bande passante en fonction de l'état de charge de chaque ST et de la disponibilité des ressources dans le réseau satellite.

Le chapitre 4 a été consacré à la caractérisation des besoins des nouveaux services destinés à être déployés dans les NGNs. Pour cela, nous avons spécifié des outils capables de générer du trafic correspondant à des sessions multimédias réelles. La plateforme de mesure implémentant l'architecture de QoS ainsi que les métriques retenues pour l'évaluer ont été ensuite définies. L'évaluation a permis de caractériser les performances de notre architecture de QoS et de valider l'adéquation de ses mécanismes vis-à-vis des contraintes des catégories d'applications envisagées tout en sauvegardant les ressources du réseau satellite.

Perspectives

Ce travail ouvre la voie à de nombreuses perspectives de recherche qui sont directement liées à l'extension de cette architecture.

En premier lieu, nous pouvons citer la possibilité d'accroître le niveau d'interaction entre les différentes couches de communication de notre architecture de QoS (approche dite *Cross Layer*) en basant par exemple le calcul des requêtes de capacité sur des informations de niveau IP voire Session. De même, des informations sur le niveau d'utilisation des ressources du terminal satellite pourront être remontées aux applications du segment utilisateur afin qu'elles puissent adapter ponctuellement en privilégiant, par exemple, des codecs générant de plus faibles débits.

Dans un souci d'interopérabilité avec les futurs réseaux NGNs, l'architecture de QoS devra également offrir la possibilité de négocier dynamiquement les SLAs aux différents clients du réseau satellite que ce soient de simples utilisateurs ou des fournisseurs d'accès Internet voire des fournisseurs de service. Tout d'abord, cela nécessite la définition d'une interface commune de négociation indépendante de la technologie de QoS employé d'un domaine administratif à un autre. Ensuite cette possibilité de négocier dynamiquement ces services entraînera notamment le besoin de provisionner dynamiquement dans le réseau satellite les ressources associées à ces services au niveau IP et au niveau MAC dans les STs. Afin d'offrir une QoS de bout en bout, les applications seront amenées à incorporer une signalisation NGN leur permettant d'effectuer cette négociation directement avec une entité satellite centrale ou avec des entités distribuées dans chaque ST qui se verront déléguer la responsabilité d'une partie des ressources du système satellite. Cette architecture impliquera le support de différentes signalisations comme par exemple :

- SIP doté de SDPs de nouvelle génération pour l'établissement de session à QoS
- COPS-SLS optimisé pour les réseaux satellites pour une négociation générique des « services supports » sur le segment satellite
- COPS-DRA (DiffServ Resource Allocation) s'appuyant sur une PIB (Policy Information Base) DiffServ étendue par des paramètres satellites de niveau MAC afin de spécifier les modifications à apporter dans les classes de service IP ou aux connexions ATM déjà établies ou à établir

Le médium satellite affichant des délais de propagation relativement long en comparaison avec les réseaux terrestres, la charge due à cette signalisation et son interactivité devra être réduite au maximum afin d'assurer le passage à l'échelle de l'architecture proposée. La minimisation des échanges de messages de signalisation, l'agrégation des réservations ou des allocations de ressources ainsi que la compression de la taille des messages sont autant de pistes à explorer afin de concrétiser cette architecture.

Par ailleurs, outre la nécessité de faire évoluer les outils, les méthodes de mesure (nouvelles métriques) et les caractérisations des applications (arrivée d'applications de téléphonie basée sur la technologie Peer-to-Peer comme Skype), notre architecture devra prendre en compte les prochaines évolutions attendues dans les réseaux satellites. Ainsi, la définition par la norme DVB-S2 de conteneurs de niveau 2 de taille variable contrairement aux conteneurs MPEG2-TS de taille fixe va permettre d'envisager de réduire les surcoûts protocolaires et de faciliter l'adoption de solutions de routage à bord du satellite.

L'approfondissement de chacune de ces perspectives permettrait de dépasser le simple problème d'intégration des réseaux satellites dans l'infrastructure NGN en les plaçant comme « la » technologie adaptée à un accès fixe ou mobile (véhiculaire notamment) aux services NGN.

Enfin, à plus long terme, l'ensemble de ces réflexions devra être élargi à un contexte « multi-réseau » où le trafic des applications sera probablement amené à traverser une série de réseaux d'accès aux caractéristiques et mécanismes de QoS fortement hétérogènes. Un parfait exemple de ce cas de figure est le couplage entre un réseau satellite et des réseaux 802.11. Chaque réseau dispose ainsi individuellement de ses propres mécanismes de QoS. Le réseau satellite s'appuie sur le DAMA et le CAC tandis que les réseaux 802.11 bénéficient des mécanismes de la norme 802.11e. Une telle étude permettrait de déterminer l'influence du couplage de ces différents mécanismes ainsi que l'intérêt d'échanger entre ces réseaux des informations sur leur état de congestion respectif ou sur les mécanismes qu'ils implémentent.

Une autre possibilité concernera l'intégration de services IP Multicast bénéficiant des propriétés naturelles de diffusion des satellites et leur interaction éventuelle avec notre architecture de QoS. De même, l'intégration de protocoles de sécurité tel qu'IPSec dans le contexte satellitaire peut également être optimisée en proposant des protocoles de distribution de clés plus performants et la mise en œuvre de tunnels IPSec plus pertinents dans le cadre de services multicast sur satellite.

Bibliographie

- [1] Skype, www.skype.com
- [2] MSN Messenger, <http://messenger.msn.com/>
- [3] ETSI/GA38(01)18: *Conclusions from the NGN Starter Group presented to ETSI GA#38*, November 2001.
- [4] ITU, "NGN 2004 Project description – version 3", published 12 February, 2004.
- [5] Claude Rigault, Rony chabine, "Cooperative computing in the control plane. Application to NGN services and control", MAN'2005, Avril 2005.
- [6] Schulzrinne H., Casner S., Frederick R., Jacobson V., "RTP: A Transport protocol for real time applications", IETF RFC 1889, janvier 1999.
- [7] ETSI TS 101 314 *Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 4; Abstract Architecture and Reference Points Definition; Network Architecture and Reference Points, V4.1.*, 2003-09.
- [8] Claude Rigault, Rony chabine, "Mécanismes coopératifs de plan contrôle global pour des services de communications multi-fournisseurs et trans-réseaux", Novembre, 2004, DNAC 2004.
- [9] A.R. Modarressi and S. Mohan, "Control and Management in Next-Generation Networks: Challenges and Opportunities", IEEE Communications Magazine, Vol. 38, No. 10, pp. 94-102.
- [10] M. Handley, H. Schulzrinne, E. Schooler, J. Rosenberg. "SIP: Session Initiation Protocol", IETF RFC 2543, 1999.
- [11] J. Rosenberg, H. Schulzrinne, G. Camarillo, A. Johnston, J. Peterson, R. Sparks, M. Handley, E. Schooler. "SIP: Session Initiation Protocol", IETF RFC 3261 2002.
- [12] M. Handley and V. Jacobson, "SDP: Session Description Protocol," IETF RFC 2327, Apr. 1998.
- [13] J. Rosenberg, et al., "An Offer/ Answer Model with the Session Description Protocol (SDP)," IETF RFC 3264, Jun. 2002.
- [14] 3GPP TS 23.228 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; IP Multimedia Subsystem (IMS); Stage 2 (Release 7), V7.0.0 (2005-06).
- [15] "Etude technique, économique et réglementaire de l'évolution vers les réseaux de nouvelle génération (NGN, Next Generation Networks)", Cabinet Arcome, Septembre 2002.
- [16] ITU-T, Recommendation Y.1291 (Y.qosar) "An architectural framework for support of Quality of Service (QoS) in Packet networks"
- [17] Shenker, S. and J. Wroclawski, "General Characterization Parameters for Integrated Service Network Elements", IETF RFC 2215, 1997.
- [18] Wroclawski, J., "The Use of RSVP with IETF Integrated Services", IETF RFC 2210, 1997.
- [19] Braden, R., et al., "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", IETF RFC 2205, 1997.
- [20] Braden, R. and L. Zhang, "Resource ReSerVation Protocol (RSVP) -- Version 1 Message Processing Rules", IETF RFC 2209, 1997.
- [21] Boyle, J., Cohen, R., Durham, D., Herzog, S., Raja, R. and A. Sastry, "COPS Usage for RSVP", IETF RFC 2749, January 2000.
- [22] P. Pan, "Scalable resource reservation signaling in the internet" Columbia University Ph.D. Thesis, 2002.
- [23] P. Pan and H. Schulzrinne, "YESSIR: A Simple Reservation Mechanism for the Internet", Int'l. Wksp. Network and Operating System Support for Digital Audio and Video (NOSSDAV), Cambridge, England, July 1998.
- [24] G. Feber et al., "Boomerang — A Simple Protocol for Resource Reservation in IP Networks", IEEE Wksp. QoS Support for Real-Time Internet Applications, Vancouver, Canada, June 1999.
- [25] Blake, Black, Carlson, Davies, Wang, Weiss, "An architecture for Differentiated Services", IETF RFC 2475, Décembre 1998.
- [26] "Methods for subjective determination of transmission quality", ITU-T Recommendation P.800, August 1996.
- [27] "The e-model, a computational model for use in transmission planning," ITU-T Recommendation G.107, March 2003.
- [28] ITU-T G.1010, "End-user Multimedia QoS categories", November 2001.
- [29] 3GPP TS 23.107, 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; "Quality of Service (QoS) concept and architecture", v6.3.0, July 2005.
- [30] ETSI TS 101 329-2: "Telecommunications and Internet Protocol Harmonization Over Networks (TIPHON) Release 3; End-to-end Quality of Service in TIPHON systems; Part 2: Definition of speech Quality of Service (QoS) classes".
- [31] BSM ETSI TS 102 295, "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM) services and architectures; BSM Traffic Classes", v1.1.1, Feb 2004.
- [32] B. Davie et al, "An Expedited Forwarding PHB", IETF RFC 3246, March 2002.
- [33] Heinanen, J., Baker, F., Weiss, W., Wroclawski, J., "Assured Forwarding PHB group", IETF RFC 2597, 1999.
- [34] A. Westerinen et al, "Terminology for Policy-Based Management", IETF RFC 3198, November 2001.
- [35] <http://www.iana.org/assignments/port-numbers>
- [36] Y. Bernet and all, "A Framework for Integrated Services Operation Over DiffServ Networks", IETF, RFC 2998, November 2000.
- [37] A. Terzis, L. Wang, J. Ogawa, L. Zhang, "A two-tier resource management model for the Internet", Proceedings of Global Internet 99, décembre 1999.
- [38] R. Yavatkar, D. Pendarakis, R. Guerin, "A Framework for Policy-based Admission Control", IETF RFC: 2753, January

- 2000.
- [39] S. Herzog, J. Boyle, R. Cohen, D. Durham, P. Rajan and A. Sastry, "COPS Usage for RSVP", IETF RFC: 2749, January 2000.
- [40] J. Boyle, R. Cohen, D. Durham, S. Herzog, R. Raja, A. Sastry, "The COPS (Common Open Policy Service) Protocol", IETF RFC 2748, January 2000.
- [41] Chan, K., Sabita, R., Hahn, S., and McCloghrie, K., "Differentiated Services Quality of Service Policy Information Base", IETF RFC 3317, March 2003.
- [42] K. Chan, D. Durham, S. Gai, S. Herzog, K. McCloghrie, F. Reichmeyer, J. Seligson, A. Smith, R. Yavatkar, "COPS Usage for Policy Provisioning", IETF RFC 3084, March 2001.
- [43] D. Goderis et al, "Attributes of a Service Level Specification (SLS) Template, <draft-tequila-sls-03.txt>, October 2003.
- [44] T.M.T. Nguyen et al, "COPS-PR Usage for SLS negotiation (COPS-SLS)", draft-nguyen-rap-cops-sls-03.txt, July 2002.
- [45] Eurescom "Inter-operator interfaces for ensuring end-to-end IP QoS: Specification of Inter-domain Quality of Service Management Interfaces", rapport de projet Eurescom P1008, mai 2001.
- [46] <http://qbone.internet2.edu/bb/bboutline2.html>.
- [47] K. Nichols, B. Carpenter, "Definition of Differentiated Services Per Domain Behaviors and Rules for their Specification", IETF RFC 3086, April 2001.
- [48] Goderis D. et al, "A service centric IP quality of service architecture for next generation networks", Proceeding of IEEE/IFIP Network Operations and Management Symposium, pp139-154, Florence, Italy, April 2002.
- [49] S. Salsano, E. Sangregorio, M. Listanti, "COPS DRA: a protocol for dynamic Diffserv Resource Allocation", Joint Planet-IP NEBULA workshop, Courmayeur, Gennaio 2002.
- [50] Qbone Signaling WorkGroup "QBone Signaling Design Team", Final Report, 2001.
- [51] S. Salsano, L. Veltri, "QoS Control by means of COPS to support SIP based applications", IEEE Networks, March/ April 2002.
- [52] L. Veltri, S. Salsano, D. Papalilo, "SIP Extensions for QoS support", IETF Draft draft-veltri-sip-qsip-01.txt, Oct 2002.
- [53] ETSI TS 101 329-5 V1.1.1 TIPHON Release 3;Technology Compliance Specification;Part 5: Quality of Service (QoS) measurement methodologies (2000-11).
- [54] J. Rosenberg, "The SIP UPDATE Method", IETF RFC 3311, October 2002.
- [55] G. Camarillo, W. Marshall, J. Rosenberg, "Integration of Resource Management and SIP", IETF RFC 3312, October 2002.
- [56] LN Hamer, B. Gage, and H. Shieh, "Framework for Session Set-up with Media Authorizatio", IETF RFC 3521, Apr. 2003.
- [57] Groupe de travail NSIS de l'IETF, <http://www.ietf.org/html.charters/nsis-charter.html>
- [58] IST EuQoS Project, End-to-end Quality of Service support over heterogeneous networks, <http://www.euqos.org>
- [59] IEEE 802.11b-1999™/Cor 1-2001, IEEE Standard for Information technology, corrigendum 1, 2001.
- [60] P.Berthou , T.Gayraud , O.Alphand , C.Prudhommeaux , M.Diaz, "A multimedia architecture for 802.11b networks", IEEE Wireless Communications and Networking Conference (WCNC'2003), New Orleans (USA), 16-20 Mars 2003, 6p.
- [61] M. Heusse, F. Rousseau, G. Berger-Sabbatel, and A. Duda. "Performance anomaly of 802.11b". In Proceedings of INFOCOM'03.
- [62] A. Doufexi, S. Armour, B.S. Lee, A. Nix and D. Bull, "An Evaluation of the Performance of IEEE 802.11a and 802.11g Wireless Local Area Networks in a Corporate Office Environment", ICC 2003, Alaska, May 2003.
- [63] Hua Zbn, Ming Li, Imrich Chlamtac, And B. Prabhakaran, "A Survey of Quality of Service In IEEE 802.11 Networks", IEEE Wireless Communications, August 2004.
- [64] S.-Y. Park et al., "Collaborative QoS Architecture between DiffServ and 802.11e Wireless LAN", Proc. IEEE VTC '03-Spring, Jeju, Korea, Apr. 2003.
- [65] M. Li et al., "End-to-end QoS Guarantee in Heterogeneous Wired-cum-Wireless Networks", Tech. rep. no. UTDCS-25-04, Univ. of TX at Dallas, July 2004.
- [66] Technical assistance in bridging the "digital divide": A Cost benefit Analysis for Broadband connectivity in Europe, PriceWaterHouseCoopers, Oct 2004.
- [67] ETSI EN 301 790 v1.3.1: "Digital Video Broadcast (DVB); Interaction channel for satellite distribution systems", ETSI Norm, January 2003.
- [68] ETSI EN 302 307 v1.1.1 "Digital Video Broadcasting, Second Generation Framing Structure, Channel Coding And Modulation Systems For Broadcasting, Interactive Services, News Gathering And Other Broadband Satellite Application".
- [69] ETSI TR 102 187 V1.1.1: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; Overview of BSM families", May 2005.
- [70] ETSI TR 101 984 V1.1.1: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; Services and Architectures", Nov 2001.
- [71] IPStar Broadband Satellite, <http://www.ipstar.com.au>
- [72] Alenia Spazio, "SkyplexNet over HB6:Network Architecture Spec.", SPE/ESW0513/ALS, July 2002.
- [73] ISO/IEC 13818-1: "Information technology — Generic coding of moving pictures and associated audio information: Systems", Second edition, December 2000.
- [74] ETSI EN 301 192 V1.4.1: "Digital Video Broadcast (DVB); DVB specification for data broadcasting", ETSI Norm,

- Nov 2004.
- [75] ETSI TR 101 202 v1.2.1: "Digital Video Broadcast (DVB); Implementation guidelines for Data Broadcasting", ETSI Technical Report, January 2003.
- [76] ISO/IEC 13818-16, "Generic coding of Moving Pictures and Associated Audio Information".
- [77] ETSI EN 300 468, Digital Video Broadcasting, "Digital Video Broadcasting: Specification for Service Information in DVB systems", 06/2004..
- [78] ISO/IEC 13818-6: "Information technology — Generic coding of moving pictures and associated audio information: Part 6: Extensions for DSM-CC", Second edition, December 2000.
- [79] IETF Internet draft "Unidirectional Lightweight Encapsulation (ULE) for transmission of IP datagrams over an MPEG-2 Transport Stream, Gorry Fairhurst, Bernhard Collini-Nocker, June 2005.
- [80] Gorry Fairhurst, Bernhard Collini-Nocker, "ULE versus MPE as an IP over DVB Encapsulation", HET-NET'04
- [81] J. Fasson, "Étude D'une Architecture Ip Intégrant Un Lien Satellite Géostationnaire", Thèse pour le doctorat en Réseaux et Télécommunications de l'Institut National Polytechnique de Toulouse, Déc. 2004.
- [82] IETF Internet draft, "Address Resolution for IP datagrams over MPEG-2 networks", Gorry Fairhurst, Bernhard Collini-Nocker, May 2005.
- [83] IETF RFC 3077, "A Link-Layer Tunneling Mechanism for Unidirectional Links", E. Duros, H. Izumiyama, N. Fujii, Y. Zhang, March 2001.
- [84] ETSI TR 101 790 v1.3.1: "Digital Video Broadcast (DVB); Interaction channel for satellite distribution systems; Guidelines for the use of EN 301 790", ETSI Technical Report, January 2003.
- [85] Grossman, D.; Heinanen, J.: "Multiprotocol Encapsulation over ATM Adaptation Layer 5", IETF RFC 2684 (obsoletes 1483), September 1999.
- [86] ETSI TR 100815 v1.1.1: "Digital Video Broadcasting (DVB); Guidelines for the handling of Asynchronous Transfer Mode (ATM) signals in DVB systems", ETSI Technical Report, February 1999.
- [87] IBIS IST Project : Integrated Broadcast Interaction System, <http://www.alcatel.es/ist-ibis>
- [88] European Space Agency, Interactive Broadband DVB-RCS/S OBP Communication System (AMERHIS), <http://www.esa.int>
- [89] DIPCAST, Projet RNRT, <http://dipcast.netvizion.fr/>
- [90] IST SATIP6 <http://satip6.tilab.com/>
- [91] QoSforRCS (AO4223) Technical Note 2 Definition of QoS Architectures, ESA AmerHis Project, May 2004
- [92] S. Dawkins, G. Montenegro, M. Kojo, V. Magret, and N. Vaidya, "End-to-end performance implications of links with errors", IETF RFC 3155: August 2001.
- [93] Border, J., Kojo, M., Griner, J., Montenegro, G., Shelby, Z., "Performance Enhancing Proxies Intended to Mitigate Link-Related Degradations", IETF RFC 3135, 2001.
- [94] ETSI TS 102 292: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia (BSM) services and architectures; Functional architecture for IP interworking with BSM networks", V1.1.1 (2004-02)
- [95] ETSI TR 101 984: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; Services and Architectures", V1.1.1 (2002-11).
- [96] ETSI TR 102 157: "Satellite Earth Stations and Systems (SES); Broadband Satellite Multimedia; IP Interworking over satellite; Performance, Availability and Quality of Service", V1.1.1 (2003-07).
- [97] Groupe SATLABS, <http://satlabs.org/>
- [98] 3GPP TR 23.802 V1.1.0, "3rd, Generation Partnership Project; Technical Specification Group Services and System Aspects; Architectural Enhancements for End-to-End Quality of Service (QoS) (Release 7)", July 2005-09-11.
- [99] Draft ETSI ES 2XX XXX: "TISPAN NGN Functional Architecture Release 1 IP Multimedia Subsystem (IMS) ", V<1.1.6> (2005-06)
- [100] IST SATIP6 project, IST-2001-3434, <http://satip6.tilab.com>
- [101] J. Border, M. Kojo, J. Griner, G. Montenegro, and Z. Shelby, "Performance enhancing proxies intended to mitigate link-related degradations", IETF RFC 3135, June 2001.
- [102] "The Flat Multicast Key Exchange protocol", IETF Draft <draft-duquer-fmke-01.txt>, L. Duquerroy, S. Josset, Sept 2004
- [103] Astra 1H, <http://www.ses-astra.com/satellites/fleet/satellites.php?sat=12>
- [104] Amerhis, <http://www.esa.int/esapub/br/br226/br226.pdf>
- [105] ETSI TR 101 790 v1.3.1: "Digital Video Broadcast (DVB); Interaction channel for satellite distribution systems; Guidelines for the use of EN 301 790", ETSI Technical Report, January 2003
- [106] Deliverable No. 1.1b : "Satellite/IPv6 Network System Scenarios", SATIP6,
- [107] Deliverable No. 2.3: "IP over Satellite Access Network functional and subsystem specification", SATIP6
- [108] K.Chan, J.Barbiaz, F.Baker, draft IETF, draft-chan-tsvmg-diffserv-class-aggr-02, July 2005
- [109] F. Delli Priscoli, A. Pietrabissa, "Design of a bandwidth-on-demand (BoD) protocol for satellite networks modeled as time-delay systems", *Automatica*, Vol. 40, Issue 5, pp. 729-741, May, 2004
- [110] Pietrabissa, A., et al., "Validation of a QoS Architecture for DVB/RCS Satellite Networks via a Demonstration Platform", to be published in *Computer Networks* 2005.
- [111] Sooriyabandara, M., Fairhurst, G. "Dynamics of TCP over BoD Satellite Networks". *International Journal of Satellite Communications: Special Issue on IP QoS for Satellite IP*, Vol. 21, pp. 427-449, July 2002.
- [112] Iera, A., Molinaro, A., Marano, S. "IP with QoS guarantees via Geo satellite channels: performance issues". *IEEE Personal*

- Communications*, Vol. 8 , N. 3 , pp.14-19, June 2001.
- [113] Mitchell, P. D., Grace, D. & Tozer, T.C. "Burst Targeted Demand Assignment Multiple-Access for Broadband Internet Service Delivery Over Geostationary Satellite". *IEEE Journal on Selected Areas in Communications*, Vol. 22, N. 3, pp. 546-558, April 2004.
- [114] Açar, G. & Rosenberg, C. "Weighted fair bandwidth-on-demand (WFBoD) for geostationary satellite networks with on-board processing". *Computer Networks*, Vol. 39, N. 1, pp. 5-20, May 15, 2002.
- [115] Deliverable No. 2.4 : "Simulation Result Report," SATIP6.
- [116] V. Firoiu, J. Kurose, and D. Towsley. "Efficient admission control for EDF schedulers". In *Proc. of IEEE INFOCOM '97*, pages 310 – 317, April 1997.
- [117] Brahms IST project (Contract IST-99-10440), <http://brabms.tilab.com>
- [118] Margouilla Runtime : <http://cqssoftware.free.fr/margouilla>
- [119] TIPHON: Definition of QoS classes, ETSI TR 101 329-2 v1.2.1, Dec. 2000
- [120] NIST-SIP-Proxy, <https://jain-sip-presence-proxy.den.java.net/>
- [121] QT, <http://www.trolltech.com/>
- [122] Linphone, <http://www.linphone.org>
- [123] Kphone, <http://www.wirlab.net/kphone/>
- [124] SIP Communicator, <https://sip-communicator.dev.java.net/>
- [125] Gnomemeeting, <http://www.gnomemeeting.org/>
- [126] Windows Messenger, <http://www.microsoft.com/windows/messenger/>
- [127] Tcpdump, <http://www.tcpdump.org>
- [128] F. Fitzek and M. Reisslein, "MPEG4 and H.263 video traces for network performance evaluation", *IEEE Network*, vol. 15, no. 6, pp. 40-54, Video traces available at <http://www.eas.asu.edu/trace>, November/December 2001.
- [129] Karagiannis, T., M. Molle, and M. Faloutsos, "Long-Range Dependence: Ten Years of Internet Traffic Modeling". *IEEE Internet Computing*, 2004.
- [130] Sabinoglu, Z. and S. Tekiney, "On Multimedia Networks: Self-Similar Traffic and Network Performance". *IEEE Communications Magazine*, 1999.
- [131] Ahmed, T., A. Mebaoua, and G. Buridant. "Encapsulation and Marking of MPEG-4 Video over IP Differentiated Services". in *ISCC'2001*. 2001.
- [132] Kimura, J., et al. "Perceived Quality and Bandwidth Characterization of Layered MPEG-2 Video Encoding". in *SPIE International Symposium on Voice, Video and Data Communications*. 1999.
- [133] Pulido, J.-M., "A Simple Admission Control Algorithm for Layered VBR MPEG-2 Streams", *Stanford University's Computer Systems Laboratory*, 2000.
- [134] Azgouna, N.B., Thèse de doctorat à l'Université Pierre et Marie Curie - Paris VI, "Etude des méthodes d'échantillonnage des flux pour la mesure dans les réseaux large bande", *Laboratoire d'Informatique de Paris VI*, 2004.
- [135] Paxson, V., "Empirically-Derived Analytic Models of Wide-Area TCP Connections". *IEEE Transactions on Networking*, 1994.
- [136] Jacobson, V., "Compressing TCP/IP Headers for Low-Speed Serial Links". *IETF RFC 1144*, 1990.
- [137] Bruce A. Mah, "An Empirical Model of HTTP Network Traffic", *Proceedings of INFOCOM'97*.
- [138] IPERF : <http://dast.nlanr.net/Projects/Iperf>

Publications

Revues internationales

- S.Combes, O.Alphand, P.Berthou, T.Gayraud. *Satellite and Next Generation Networks: QoS Issues*, Novembre 2004, 29p, à paraître dans International Journal of Space Communications 2005.
- A.Pietrabissa, T.Inzerilli, O.Alphand, P.Berthou, E.Fromentin, T.Gayraud, F.Lucas. *Validation of a QoS architecture for DVB/RCS satellite networks via a hardware demonstration platform*, soumis en Juillet 2004, 23p, accepté pour publication dans Computer Networks 2005 (Elsevier, Grande-Bretagne).

Conférences internationales

- O. Alphand, P. Berthou, T. Gayraud, S. Combes, QoS Architecture over DVB-RCS satellite networks in a NGN framework, Globecom 2005
- S.Combes, C.Baudoin, N.Sannier, G.Smaoui, O.Alphand, P.Berthou, T. Gayraud, QoS Architecture for DVB-RCS Networks, 11th Ka and Broadband Communications Conference and 23rd AIAA International Communications Satellite Systems Conference (ICSSC), 25-28 Septembre 2005, 8p
- O.Alphand, P.Berthou, *Dynamic QoS session set-ups for next generation satellite networks*, IEEE Infocom 2005, Student Workshop, Miami, FL, USA, 14 Mars 2005, 2p.
- O.Alphand, P.Berthou, T.Gayraud, S.Josset, E.Fromentin. *SATIP6: Next Generation Satellite System Demonstrator*, IFIP 18th World Computer Congress. Broadband Satellite Communication Systems (BSCS'04), Toulouse (France), 22-27 Aout 2004, pp.35-44.
- O.Alphand, P.Berthou, T.Gayraud, S.Josset, E.Fromentin. *SATIP6: Satellite Testbed for Next Generation Protocols*, 2nd International Working Conference on Performance Modelling and Evaluation of Heterogeneous Networks (HET-NETs'04), Ilkley (GB), 26-28 Juillet 2004, pp.P45/1-P45/11.
- L.Duquerroy, S.Josset, O.Alphand, P.Berthou, T.Gayraud. *SatIPSec: an optimized solution for securing multicast and unicast satellite transmissions*, 22nd AIAA International Communications Satellite Systems Conference (ICSSC), Monterey (USA), 9-12 Mai 2004, 11p.
- P.Berthou, T.Gayraud, O.Alphand, C.Prudhommeaux, M.Diaz. *A Multimedia Architecture for 802.11b Networks*, IEEE Wireless Communications and Networking Conference (WCNC'2003), New Orleans (USA), 16-20 Mars 2003, 6p.

Conférences nationales

- O.Alphand, P.Berthou, T.Gayraud, S.Combes. *Signalisation de QoS dans un réseau satellite de nouvelle génération*, 11eme Colloque Francophone sur l'Ingénierie des Protocoles (CFIP'2005), Bordeaux (France), 29 Mars - 1er Avril 2005, pp.483-498.
- O.Alphand, P.Berthou, T.Gayraud, M.Diaz. *Une architecture multimedia pour les réseaux sans fil 802.11b*, Colloque de l'Ecole Doctorale Informatique et Télécommunications (EDIT'03), Toulouse (France), 14-15 Avril 2003, pp.92-97.

Rapports de contrat et rapports techniques

- O.Alphand, P.Berthou, T.Gayraud, S.Combes, A.Pietrabissa, T.Inzerilli. A QoS architecture for DVB-RCS next generation satellite networks, Rapport LAAS N°05005, Janvier 2005, 13p,
- O.Alphand. *Evaluation de la qualité de service de PLATINE*, Rapport LAAS N°04750 Contrat ALCATEL N° A76804, Decembre 2004, 27p.

- O.Alphand , P.Berthou , T.Gayraud , E.Fromentin , F.Lucas , S.Josset. *IPv6 satellite testbed implementation and validation: system validation plan*, Rapport LAAS N° 04267 SATIP6, Project IST-2001-34344, Mai 2004, 13p.
- O.Alphand , P.Berthou , T.Gayraud , E.Fromentin , F.Lucas , L.Duquerroy , S.Josset , J.Kuhnle , I.Melhus. *Trial measurement results & evaluation report*, Rapport LAAS N°04236 SATIP6, Project IST-2001-34344, Avril 2004, 82p.
- T.Gayraud , D.Raymond , P.Berthou , O.Alphand , M.Diaz , S.Josset , S.Combes, *Evaluation of multimedia applications over broadband satellite systems*, Rapport LAAS N°03539, Decembre 2003, 4p.