

UNIVERSITÉ DE BORDEAUX I
ÉCOLE DOCTORALE DE MATHÉMATIQUES ET
D'INFORMATIQUE

Mémoire présenté en vue de l'obtention de l'
HABILITATION A DIRIGER DES RECHERCHES

Discipline : **INFORMATIQUE**

Présenté par
Toufik AHMED

Transport Adaptatif et Contrôle de la Qualité des Services
Vidéo sur les Réseaux IP Filaires, Sans-fil et sur les
Architectures P2P

Soutenue le JJ/mois/2008 devant le Jury composé de :

M.	Ahmed Karmouch	Professeur, Université d'Ottawa	Rapporteur
M	George Pavlou	Professeur, University College London.	Rapporteur
M.	Olivier Festor	Directeur de recherche, LORIA, Nancy	Rapporteur
M.	Raouf Boutaba	Professeur, University of Waterloo	Examineur
M.	Abdelmadjid Bouabdellah	Professeur, Université de Technologie de Compiègne	Examineur
Mme.	Francine Krief	Professeur, ENSEIRB, Bordeaux	Examineur
M.	Richard Castanet	Professeur, ENSEIRB, Bordeaux	Examineur

You are not allowed to view this page

One drop in the ocean of thought

Transport Adaptatif et Contrôle de la Qualité des Services Vidéo sur les Réseaux IP Filaires, Sans-fil et sur les Architectures P2P

RESUME

De nouveaux services de distribution de contenus audiovisuels et multimédia ont fait leur apparition sur les réseaux fixes (ADSL, câble, ...) et mobile (UMTS, WiFi, DVB-H/T,...). Ces dernières années, afin de faciliter les échanges de fichiers et de contenus multimédia, des réseaux P2P se sont aussi développés au dessus des réseaux classiques. Cependant, plusieurs caractéristiques de ces réseaux font que le déploiement à grande échelle de services temps-réel présente actuellement un défi par rapport à des applications telles que le courrier électronique (email) et le Web. En effet, le protocole IP est devenu le mode de transport de référence sur ces différents réseaux, ce qui pose certains problèmes liés à la qualité de service (QoS) hérités du modèle protocolaire IP. Les applications audiovisuelles doivent faire face aux différentes variations de la bande passante et du délai de transfert, tout en restant robustes face aux pertes de paquets et aux erreurs de transmission. C'est dans ce cadre que s'inscrivent mes activités de recherche.

Mes travaux ont porté sur l'analyse des performances d'algorithmes et de protocoles dédiés au transport adaptatif et au contrôle de qualité des services de streaming vidéo temps-réel sur les réseaux IP : *filaires, sans-fil* et sur les architectures *P2P*. L'originalité de mes travaux repose sur la proposition d'un système adaptatif de transport vidéo respectant l'approche intégrée, également appelée « cross-layer » pour les réseaux filaires fixes et sans-fil de type IEEE 802.11. En effet, tandis que la plupart des travaux de recherche dans ce domaine respecte le paradigme d'isolation et d'indépendance des couches protocolaires héritées du modèle de référence ISO, ma proposition conceptuelle consiste à supprimer cette limite en autorisant une meilleure prise en charge de la sémantique et des contraintes applicatives au niveau des couches et des mécanismes dédiés au transport.

Mes contributions s'articulent autour de trois principaux axes de recherches. Premièrement, un système de transport adaptatif de flux MPEG-4 vidéo sur un réseau filaire est présenté. Celui-ci intègre des mécanismes de classification automatique d'objets audiovisuels couplés à des algorithmes de différenciation de service et d'ajustement de débit. Deuxièmement, et pour améliorer les performances et la fiabilité des services vidéo transportés sur les réseaux sans fil 802.11 dont l'état du canal varie considérablement comparé à un canal filaire, plusieurs mécanismes d'interaction entre couches ont été proposés, on peut citer par exemple le contrôle des erreurs, la fragmentation et le groupement des images. Finalement, des mécanismes de contrôle de QoS pour le transport multi-sources des flux vidéo sur les réseaux P2P sont proposés tels que des mécanismes de sélection des meilleurs pairs participant à la session, des mécanismes de commutation de flux et de pairs, et des mécanismes de conception de réseaux virtuels overlays.

Mots-clés : Réseaux IP, LAN sans-fil, P2P, transport vidéo temps réel, contrôle d'erreurs, contrôle de flux, qualité de service.

Adaptive Streaming and Control of Video Quality of Service over Wired/ Wireless IP Networks and P2P Architectures

ABSTRACT

There is an increasing demand of delivering new multimedia services over different networks types including wired and wireless networks (ADSL, Cable, UMTS, WiFi, DVB-H/T ...). At the same time, P2P networks have also been developed on top of classical IP networks to ease the exchanges of files and media content. However, various network characteristics make the large-scale deployment of multimedia services more challenging than traditional Internet applications such as email and Web. These services that transmit real-time audiovisual data over IP must cope with the time varying bandwidth and delay of the network while being resilient to packet loss and errors. My research work has been carried out in this framework.

The research work presented in this document aims to analyse and to develop new protocols and mechanisms for adaptive streaming and control of video quality of service over wired/wireless IP networks and P2P architectures. The originality of this work is to propose a cross-layer approach for resolving some of the critical issues on delivering packet video over wired and wireless 802.11 IP networks with satisfactory quality of service. While, current and past works on this topic respect the protocol layer isolation paradigm, the key idea behind our work is to break this limitation and to rather inject content-level semantic and service-level requirements within the proposed IP video transport mechanisms and protocols.

My contributions hinges on three important research area. First, an efficient and adaptive cross layer video delivery system over IP network has been proposed. This system is composed of a system-level audiovisual object classification model augmented with a dynamic network-level packet video marking algorithm, and a robust and adaptive application level framing protocol to allow fine-grained TCP-friendly rate control and unequal error protection. Second, to enhance the reliability of the transmitted packet video over 802.11 wireless networks, in which the canal condition changes drastically, different mechanisms have been proposed such as joint error/throughput control, MAC-level frame fragmentation, and video frame grouping. Finally, different mechanisms for controlling the QoS of multi-sources video streams over P2P network have been proposed including best peers selections, peer/stream switching, and overlay network organization.

Key-words: IP network, wireless Local Area Network, P2P, real-time video streaming, errors control, quality of services.

Table des Matières

1	INTRODUCTION GENERALE.....	10
1.1	CONTEXTE DE RECHERCHE	10
1.2	ORGANISATION DU DOCUMENT.....	13
2	CURRICULUM VITAE	14
2.1	NOTICE INDIVIDUELLE.....	14
2.2	DIPLOMES	14
2.3	EXPERIENCE PROFESSIONNELLE	15
2.4	SEJOURS/VISITES DANS DES LABORATOIRES DE RECHERCHE	15
2.5	ANIMATIONS SCIENTIFIQUES.....	16
2.6	SEMINAIRES, TUTORIAUX ET ARTICLES INVITES.....	17
2.7	THEMATIQUES DE RECHERCHE.....	18
2.7.1	<i>Transport et Contrôle de la qualité des services vidéo sur réseaux IP filaires.....</i>	<i>19</i>
2.7.2	<i>Transport et Contrôle de la qualité des services vidéo sur les réseaux sans fil 802.11</i>	<i>20</i>
2.7.3	<i>Transport et Contrôle de la qualité des services vidéo sur les réseaux P2P.....</i>	<i>21</i>
2.8	RESPONSABILITES ADMINISTRATIVES ET SCIENTIFIQUES	22
2.9	COLLABORATIONS NATIONALES ET INTERNATIONALES	22
2.10	FINANCEMENT ET CONTRATS DE RECHERCHE	23
2.11	JURY DE THESES ET ENCADREMENTS D'ÉTUDIANTS.....	23
2.12	ACTIVITES D'ENSEIGNEMENT	24
2.13	LISTE DES PUBLICATIONS	24
2.13.1	<i>Dans des Editions de Livres</i>	<i>24</i>
2.13.2	<i>Dans des Chapitres de Livres.....</i>	<i>25</i>
2.13.3	<i>Dans des Revues Internationales (avec comité de lecture)</i>	<i>25</i>
2.13.4	<i>Dans des Conférences Internationales (avec Comité de Lecture).....</i>	<i>26</i>
2.13.5	<i>Dans des Conférences Nationales (avec Comité de Lecture).....</i>	<i>29</i>
2.13.6	<i>Autres publications.....</i>	<i>29</i>
3	TRANSPORT ET CONTROLE DE LA QUALITE DES SERVICES VIDEO SUR IP.....	30
3.1	INTRODUCTION	30
3.2	ÉTAT DE L'ART	31
3.2.1	<i>Gestion de la bande passante.....</i>	<i>31</i>
3.2.2	<i>Gestion des pertes et des erreurs</i>	<i>35</i>
3.2.3	<i>Gestion du délai de bout-en-bout et de la gigue.....</i>	<i>38</i>
3.2.4	<i>Gestion des services audiovisuels : les normes ISO MPEG-4, MPEG-7 et MPEG-21</i>	<i>40</i>
3.2.5	<i>Les profils de transport temps réel (RTP) pour les services vidéo MPEG4.....</i>	<i>44</i>
3.3	CONTRIBUTIONS	47
3.3.1	<i>Un modèle de Classification automatique d'Objets Audiovisuels MPEG-4</i>	<i>47</i>
3.3.2	<i>RTP4mux : Un profil de transport RTP pour le multiplexage de flux MPEG-4.....</i>	<i>50</i>
3.3.3	<i>Un Algorithme de contrôle du débit « TCP-friendly » pour les sources vidéo MPEG-4</i>	<i>52</i>
3.4	CONCLUSION	55
3.5	MES PUBLICATIONS ASSOCIEES.....	55
3.6	BIBLIOGRAPHIE	56
4	TRANSPORT ET CONTROLE DE LA QUALITE DES SERVICES VIDEO SUR RESEAUX SANS-FIL	59
4.1	INTRODUCTION	59
4.2	ÉTAT DE L'ART	59
4.2.1	<i>Introduction au concept Cross-Layer (XL)</i>	<i>62</i>
4.2.2	<i>Signalisation dans les architectures Cross-Layer</i>	<i>63</i>
4.2.3	<i>Les interactions Cross-Layer</i>	<i>64</i>

4.2.4	Les projets européens traitant la problématique Cross-Layer	65
4.2.5	Les réseaux WLAN 802.11	66
4.3	CONTRIBUTIONS	67
4.3.1	Adaptation du débit vidéo en fonction du débit physique 802.11	68
4.3.2	Adaptation conjointe FEC/débit vidéo en fonction de la puissance du signal et des taux de pertes....	70
4.3.3	Fragmentation adaptative cross-layer au niveau MAC 802.11	72
4.3.4	Groupeage de trames basé sur une image vidéo.....	74
4.4	CONCLUSION	76
4.5	MES PUBLICATIONS ASSOCIEES	77
4.6	BIBLIOGRAPHIE	77
5	TRANSPORT ET CONTROLE DE LA QUALITE DES SERVICES VIDEO SUR LES RESEAUX P2P	79
5.1	INTRODUCTION	79
5.2	ÉTAT DE L'ART	80
5.2.1	Taxonomie des réseaux P2P.....	80
5.2.2	Applications utilisant les réseaux P2P.....	82
5.2.3	Les systèmes de streaming multimédia sur P2P.....	83
5.2.4	Problèmes de streaming vidéo sur P2P.....	85
5.3	CONTRIBUTIONS	86
5.3.1	Mécanisme de sélection des pairs	88
5.3.2	Mécanisme d'adaptation vidéo sur P2P.....	90
5.3.3	Concept de collaboration P2P/Fournisseur.....	94
5.4	CONCLUSION	95
5.5	MES PUBLICATIONS ASSOCIEES	96
5.6	BIBLIOGRAPHIE	96
6	CONCLUSION GENERALE ET PERSPECTIVES DE RECHERCHE	99
6.1	CONCLUSION GENERALE	99
6.2	PERSPECTIVES DE RECHERCHE	100
6.2.1	La mobilité	101
6.2.2	Les nouveaux modes de distribution de contenus multimédia.....	102
6.2.3	Les réseaux de capteurs	103
6.2.4	Les futurs réseaux P2P collaboratifs	103
6.3	BIBLIOGRAPHIE	104
	ANNEXE : ARTICLES SELECTIONNES	105

Table des Figures

FIGURE 3.1 : PRINCIPALES NORMES DE COMPRESSION AUDIOVISUELLES ISSUES DE L'ISO ET DE L'ITU.	32
FIGURE 3.2 : TECHNIQUES DE CONTROLE D'ERREURS ET DE PERTES.	36
FIGURE 3.3 : COMPROMIS DELAI / PERTES POUR UN SERVICE MULTIMEDIA INTERACTIF.	40
FIGURE 3.4 : STRUCTURE D'UNE SCENE MPEG-4 ET LES DESCRIPTEURS D'OBJETS	41
FIGURE 3.5 : LES ELEMENTS ARCHITECTURAUX DE MPEG-21.....	43
FIGURE 3.6 : MACHINE D'ADAPTATION DANS MPEG-21.....	44
FIGURE 3.7 : TRANSPORT DES FLUX MPEG-4 AVEC UN EN-TETE SL REDUIT	45
FIGURE 3.8 : REPRESENTATION DES COUCHES DANS H.264 SVC (SCALABLE VIDEO CODING).....	48
FIGURE 3.9 : ARCHITECTURE DU CLASSIFICATEUR RBF	49
FIGURE 3.10 : ILLUSTRATION DU PROCESSUS D'ENCAPSULATION DANS RTP4MUX	51
FIGURE 4.1 : L'ADAPTATION CONJOINTE DU TAUX DE REDONDANCE FEC ET DU DEBIT VIDEO	71
FIGURE 4.2 : LES DIFFERENTS ETATS DU SYSTEME FEC ADAPTATIF	72
FIGURE 4.3 : PRINCIPE DE GROUPEMENT DE TRAMES.....	75
FIGURE 4.4 : GROUPEMENT DE TRAMES MAC A BASE D'IMAGES VIDEO	76
FIGURE 5.1 : STREAMING DE LA VIDEO SVC MULTI-SOURCES SUR UN RESEAU P2P.....	87
FIGURE 5.2 : ORGANISATION HYBRIDE DU RESEAU P2P POUR LE STREAMING VIDEO	90
FIGURE 5.3 : MODELES D'ADAPTATION DES FLUX SVC	91
FIGURE 5.4 : ORGANISATION HYBRIDE DU RESEAU P2P POUR LE STREAMING.....	95

Liste des Tableaux

TABEAU 3.1 : PROBLEMES ET SOLUTIONS LIES A LA DISTRIBUTION DES SERVICES AUDIOVISUELS SUR IP.....	31
TABEAU 3.2 : COMPLEXITE ET DOMAINE D'APPLICATION DES TECHNIQUES DE GESTION DE LA BANDE PASSANTE.	35
TABEAU 3.3 : COMPLEXITE ET DOMAINE D'APPLICATION DES TECHNIQUES DE GESTION DES PERTES/DES ERREURS.	38
TABEAU 3.4 : LES DESCRIPTEURS MPEG-7 DE QUALITE DE SERVICE ASSOCIES AUX AVO.....	47
TABEAU 3.5 : EXEMPLE D'OVERHEAD INTRODUIT PAR LE TRANSPORT DE FLUX MPEG-4 A BAS DEBIT	50
TABEAU 5.1 : CLASSIFICATION DES RESEAUX P2P SELON LEUR DOMAINE D'APPLICATION.....	83
TABEAU 5.2 : EXEMPLE DE NOMBRE DE NŒUDS EN FONCTION DE LA POPULARITE DU CONTENU	89
TABEAU 5.3 : PARAMETRES UTILISES DANS LA DESCRIPTION DU TRAFIC	91

1 Introduction Générale

1.1 Contexte de recherche

Les réseaux de nouvelle génération sont devenus de plus en plus complexes, interconnectant différentes technologies et architectures (IP, WiFi, DVB-T/S, UMTS, GSM/GPRS) à travers plusieurs plates-formes matérielles et logicielles pour offrir une multitude de services (données, téléphonie, diffusion vidéo, vidéoconférence, etc.). Ces nouvelles technologies réseaux ont permis une augmentation spectaculaire des capacités, notamment à l'accès. Ceci-ci a suscité un grand intérêt pour la distribution massive de services audiovisuels et multimédia temps réel. Cependant, la garantie de la qualité de service (QoS) de bout-en-bout sur ces réseaux hétérogènes pour ces types de service pose actuellement un défi majeur à l'ensemble des acteurs de ce domaine : industriels, universitaires, instituts de recherche, consortiums et instituts de normalisation. Quelques solutions ont été proposées mais celle-ci restent partielles puisqu'elles traitent de cas très particuliers et applicables à certains réseaux, services et terminaux. Chaque acteur, expert dans son domaine, tente d'imposer sa vision et ses solutions qui sont étroitement associées à sa culture scientifique, pour apporter son savoir-faire dans la mise en place de la convergence des réseaux et des services.

En parallèle, plusieurs avancées technologiques dans le domaine de la compression audio/vidéo à savoir MPEG-4 part 2, MPEG-4 part 10 ou H.264/AVC (Advanced Video Coding), et récemment son extension SVC (Scalable Video Coding), vont incontestablement favoriser la convergence des services audiovisuels et multimédia vers IP.

De nos jours, le marché audio-visuel et multimédia tend vers la convergence d'infrastructures de réseaux et l'intégration des technologies fixes, mobiles et de diffusion, afin de créer un nouvel environnement au sein duquel les utilisateurs pourront accéder à tout type de service audio-visuel et multimédia n'importe où et n'importe quand et en utilisant divers équipements. Ceci représente aujourd'hui un enjeu majeur et un levier de croissance. Le défi actuel est de proposer aux utilisateurs finaux une grande variété de services (programmes audio-visuels, Internet, applications multimédia, téléphonie, etc.) au travers d'infrastructures convergentes fiables et qui permettent la prise en charge de la QoS.

La convergence de ces différents secteurs est souhaitable et inéluctable mais soulèvent de nombreuses interrogations et problématiques techniques et scientifiques. Dans ce contexte, un nombre important de réseaux de communications, tels que le WLAN, WMAN et cellulaires 2G/3G (GSM, GPRS, UMTS), peuvent être considérés comme candidats pour la distribution des services audiovisuels et multimédia. Réciproquement, les réseaux de télédiffusions numériques (DVB), satellite (DVB-S), terrestre (DVB-T) et pour terminaux mobiles

(DVB-H) peuvent être améliorés en se basant sur IP comme moyen de transport pour offrir l'interactivité des services et se positionner ainsi au cœur de la convergence. Il faut noter que cette dernière décennie a vu se concrétiser la convergence des services de télécommunications classique et les services IP et plus généralement de l'Internet, notamment, grâce à la démocratisation des offres de services des « triple play » (Internet, téléphonie par IP, TV par IP). Ce phénomène est, actuellement, en pleine effervescence dans le domaine mobile, du fait de l'adoption d'une nouvelle architecture standardisée et ouverte pour la gestion de services appelée IP Multimedia sub-System (ETSI/3GPP IMS). Le protocole IP joue le rôle fédérateur en offrant une connectivité de bout-en-bout indépendamment de la nature des réseaux sous-jacents et de la couche d'accès ou encore du support utilisé. Ce concept du « Tout-IP » ou « All-IP » constitue une évolution majeure pour les opérateurs (réseaux de télécommunications fixes ou mobiles, opérateurs de services, opérateurs virtuels, etc.) dans la mesure où ils vont pouvoir déployer et offrir des services à valeur ajoutée et à moindre coût. Le concept introduit par l'IMS consiste à concevoir une architecture centrée autour des services utilisateurs. Il recommande la mise en œuvre d'une architecture de services horizontale et « en couches ou plans » afin de pouvoir réutiliser les fonctions de contrôle existantes et de développer très rapidement de nouvelles applications/services indépendamment du réseau utilisé. Le principal gain pour les opérateurs réside dans la rapidité de pérennisation des investissements consentis dans cette nouvelle architecture réseau. L'IMS offre un socle de déploiement à grande échelle de nouveaux services et d'applications, tels que la messagerie vocale, la redirection d'appel, les serveurs vocaux interactifs, le push-to-talk, le partage de bureau, le tableau blanc, etc. Ce qui constitue une opportunité pour les opérateurs de services pour créer et pénétrer d'autres parts de marché numérique et audiovisuel. Les utilisateurs pourront avoir accès à de nouveaux services, des chaînes de télévision aux contenus innovants, des services TV linéaires ou non linéaires, de type pay-per-view interactif, et au format adapté à la taille des écrans utilisés et aux préférences des utilisateurs. Cependant, l'infrastructure réseaux doit suivre cette tendance en offrant des mécanismes qui permettent de mieux contrôler et de prendre en charge la qualité de service.

Il est clair que le problème de la qualité de service perçue par l'utilisateur final est au-delà de la QoS purement réseau, puisque elle inclut plusieurs dimensions telles que la sécurité, la disponibilité et, principalement, la capacité du réseau à acheminer et à adapter le service pour satisfaire l'utilisateur final. Afin de répondre au besoin de l'utilisateur en terme de QoS et plus précisément de QoE (Quality of Experience), il est nécessaire de concevoir des architectures et des protocoles réseaux capables de tenir compte, à la fois, du contenu transporté (profil du contenu), des caractéristiques du terminal (profil du terminal), du profil des clients et des mécanismes de QoS offerts dans le réseau sous-jacent (profil réseau).

La plupart des travaux de recherche existant ou mené dans ce domaine respecte le paradigme d'isolation et d'indépendance des couches protocolaires hérité du modèle de référence des réseaux ISO/OSI (International Standard Organization / Open System Interconnection). **Mon approche conceptuelle supprime cette limite**

en autorisant une meilleure prise en charge de la sémantique et des contraintes applicatives (QoS au niveau applicative) par les couches et les mécanismes dédiés au transport (QoS réseaux) et vis-versa.

Le concept proposé est de supporter la continuité de la QoS horizontalement (à travers les acteurs de la chaîne de service) et verticalement (dans la pile protocolaire de transport) pour offrir un service de bout-en-bout.

Il est clair que dans le contexte d'un réseau à infrastructures convergentes et multiservices, les techniques de garantie de QoS par des mécanismes de réservation, de planification et de dimensionnements traditionnels sont mal adaptées et permettent pas le passage à l'échelle (scalability). En effet, la nature hautement sporadique et hétérogène des flux transportés, comme les variations rapides des quantités de trafic à transporter, nécessitent la mise en place de mécanismes de contrôle plus dynamiques aussi bien au niveau du réseau qu'au niveau des terminaux consommant le flux. Le défi majeur est de prédire correctement le trafic et donc le comportement du réseau pour mettre en place la solution de QoS. Cela se traduit par le fait qu'il n'est pas possible de contrôler le réseau de manière statique en boucle ouverte, mais requiert plutôt d'observer l'état du réseau pour décider des stratégies de reconfiguration et d'adaptation nécessaires aussi bien au niveau des mécanismes de contrôle de QoS, au niveau réseaux ou encore au niveau du flux transporté.

À cet égard, le niveau de QoS demandé initialement par un utilisateur final doit être maintenu partout dans la chaîne de transport dans une variété de points, depuis la génération du contenu jusqu'à sa consommation à travers plusieurs réseaux. Le cas échant, le niveau de QoS doit être adapté à la quantité des ressources disponibles tout en garantissant le contrat de service. D'après ce concept, le niveau de QoS est mesuré dans multiples points au niveau du terminal final (l'utilisateur), à l'entrée et à la sortie de chaque réseau ainsi qu'à la génération du service.

Ces améliorations constituent un défi pour les services multimédia qui doivent prendre en compte diverses caractéristiques au niveau applicatif, être en mesure d'optimiser les ressources utilisées et assurer une meilleure qualité perçue pour l'utilisateur final. Par ailleurs, l'hétérogénéité présente à tous les niveaux (réseaux, terminaux, contenu et systèmes), constitue un point important qui ralentit la convergence globale des réseaux et des services. L'interopérabilité entre systèmes hétérogènes est inévitable pour le déploiement à grande échelle des services.

C'est dans ce cadre de double convergence des services voix/vidéo et données, et convergence des réseaux fixes, mobiles et broadcast que s'inscrit mon activité de recherche depuis 1999. En mettant l'accent sur la conception, l'implémentation et l'évaluation des performances d'algorithmes et de protocoles de communications pour le transport de flux vidéo sur les réseaux IP filaires, sans-fil et P2P. Plus particulièrement, mes travaux de recherche, ont tout d'abord, porté sur le transport des flux vidéo MPEG-4 sur les réseaux IP avec différenciation de services, puis sur la proposition de mécanismes « cross-layer » de contrôle de QoS des flux vidéo transportés sur les réseaux sans-fil 802.11, et finalement sur le transport des flux vidéo sur les réseaux P2P.

1.2 Organisation du document

Ce document est organisé en 5 chapitres et une annexe.

Le chapitre 2 présente mon curriculum vitae, puis une synthèse de mes activités de recherche et d'enseignement, de mes responsabilités administratives et scientifiques, de mes participations à des projets, d'encadrement d'étudiants ainsi que de la liste de mes publications.

Le chapitre 3 décrit les problématiques et mes travaux de recherche sur le « transport et le contrôle de la qualité des services vidéo sur les réseaux IP ».

Le chapitre 4 décrit les problématiques et mes travaux de recherche concernant le « transport et le contrôle de la qualité des services vidéo sur les réseaux sans-fil 802.11 ».

Le chapitre 5 décrit les problématiques et mes travaux de recherche pour le « transport et le contrôle de la qualité des services vidéo sur les réseaux P2P ».

Le chapitre 6 conclut ce document et présente certaines perspectives de recherche.

Enfin, l'annexe reproduit quelques unes de mes publications.

2 Curriculum Vitae

2.1 Notice Individuelle

Nom patronymique	AHMED
Prénom	Toufik
Date et Lieu de naissance	03 Août 1976 à Blida, Algérie.
Nationalité	Algérienne
Situation familiale	Célibataire
Grade	Maître de conférences - classe normale
Etablissement	E.N.S.E.I.R.B (Ecole Nationale Supérieure d'Electronique, Informatique et de Radiocommunications de Bordeaux)
Section CNU	27
Email	@ tad@labri.fr
Web	 http://www.labri.fr/~tad

Adresse professionnelle

Laboratoire LaBRI
Université Bordeaux 1
351 cours de la Libération
33402 – Talence

Adresse personnelle

Résidence Bagatelle
Bat. B2, Apt. 49
6 rue Jean Zubietta
33400 - Talence

2.2 Diplômes

Nov. 2003 **Doctorat** de l'université de Versailles Saint Quentin-En-Yvelines en Informatique, titulaire d'une bourse MESR, mention très **honorable avec les félicitations du jury**.

Intitulé : **Diffusion Adaptative des Paquets Vidéo sur IP: Une Approche Cognitive**

Lieu : Equipe Réseaux, laboratoire PRISM, Université de Versailles

Soutenue publiquement le : 25 Nov. 2003

Manuscrit disponible à l'adresse: <http://www.labri.fr/~tad/thesis.pdf> (2.1 Mo)

Jury :

Samir	Thomé	ENST, Paris	Président
Ahmed	Karmouch	Univ. Ottawa	Rapporteur
Eric	Horlait	Univ. Paris VI	Rapporteur
Jean-Pierre	Claudé	Univ. Versailles	Directeur de thèse
Ahmed	Mehaoua	Univ. Versailles	Directeur de thèse
Raouf	Boutaba	Univ. Waterloo	Examineur
Olimpiu	Negru	Thales TBM	Examineur

Sep. 2000 DEA MISI (Méthodes Informatiques des Systèmes Industriels), option réseaux, mention bien (**Major de la promotion**), Université de Versailles Saint Quentin-En-Yvelines.

Responsable du DEA : Prof. Claude Timsit

Oct. 1999 Ingénieur d'état en Informatique (Systèmes Informatiques), mention très bien (**Major de la promotion**), Institut National d'Informatique, INI, Alger.

Juin 1994 Baccalauréat Sciences, mention bien. Blida, Algérie

2.3 Expérience Professionnelle

Depuis Sep. 2004 - Maître de Conférences (Enseignant / Chercheur)

- **Etablissement** : Ecole Nationale Supérieure d'Electronique, Informatique et de Radiocommunications de Bordeaux ([E.N.S.E.I.R.B.](#)), (Directeur Prof. Richard Castanet)

- **Recherche** : Laboratoire de Recherche en Informatique de Bordeaux : [CNRS LaBRI](#), UMR 5800, (Directrice Prof. Maylis Delest).

- **Equipe de recherche** : [COMET](#) (Responsable : Prof. Francine Krief)

Sep. 2003- Août. 2004 Attaché Temporaire d'Enseignement et de Recherche (**ATER**) en Informatique avec une demi charge d'enseignement (92h), Université de Versailles Saint Quentin-En-Yvelines.

2.4 Séjours/Visites dans des laboratoires de recherche

Fév. 2002 – Août 2002 Université de Waterloo, Canada

(7 mois)

Département d'Informatique

Equipe du professeur : Raouf Boutaba

Juillet 2004

(2 semaines)

NCSR Demokritos, Institute of Informatics & Telecommunications (IIT), Athènes, Grèce.

Equipe du professeur : Anastasios Kourtis

2.5 Animations scientifiques

Membre du comité technique des congrès TPC

- IEEE International Conference on Communications (ICC):
 - Wireless Networking Symposium pour ICC2005 WN, ICC 2006 WN, ICC 2007 WN, ICC 2008 WN,
 - ICC'09 SAC (Selected Area in Communication Symposium), ICC'09 WN (Wireless Networking Symposium), ICC'09 CISS (Communication and Information Systems Security Symposium)
- IEEE Globecom
 - Ad Hoc, Sensor and Mesh Networking Symposium (AHSN) 2007/2008 :
 - Computer and Communications Network Security Symposium2008
 - World Class Solutions: Networking the Globe Symposium 2006
 - Multimedia Communications Symposium 2006
- ICC Workshop CAMAD 2008 (13th IEEE International Workshop on Computer Aided Modeling, Analysis and Design of Communication Links and Networks)
- 12th International Workshop on Computer Aided Modeling, Analysis and Design of Communication Links and Networks (CAMAD 2007)
- Second Workshop on MultiMedia Applications over Wireless Networks (MediaWiN 2007) within ISCC 2007
- IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) 2007, 2008
- Next Generation Networking Middleware' NGNM05

Co Président de conférence

- Co Président de la conférence internationale MMNS (11th IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services) / MANWEEK 2008 - 22-23 September 2008.
Web: <http://magellan.tssg.org/2008/mmns/cfp.php?lang=en>
- Co-chairs:
 - George Pavlou (University College London, United Kingdom)
 - Toufik Ahmed (CNRS LaBRI, University Bordeaux 1, France)
 - Tasos Dagiuklas (TEI of Mesolonghi, Greece)

Co Président de workshop/Session

- Co Président de Session Special (Workshop) ADVISE 2008 “Advanced Video Services Delivery over Heterogeneous Networks”, in conjunction with 9th International, Workshop on Image Analysis for Multimedia Interactive Services, May 7-9, 2008, Klagenfurt, Austria,
Web: <http://www.wiamis2008.org/specialSessionADVISE.html>
 - Co-chairs:
 - Toufik Ahmed (CNRS LaBRI Lab. – University Bordeaux 1, France),
 - Ingo Kofler (Department of Information Technology (ITEC), Klagenfurt University, Austria),
 - Olimpiu Negru (Thomson Broadcast & Multimedia, France)
- Co Président du workshop AMIS-NET 2008 “Adaptive Multimedia and IPTV Streaming over P2P Networks”, in International Conference on Telecommunications and Multimedia (TEMU 2008)
Web: <http://www.temu.gr/2008/filez/AMIS-NET-2008.pdf>
 - Co-chairs:
 - Toufik AHMED, CNRS LaBRI Lab. - University Bordeaux 1
 - Charalabos SKIANIS, University of the Aegean, Greece
- Président de la session de présentation « Session 2 : Content Protection, Delivery,

and Terminal) dans le workshop “End-to-End QoS through Integrated Management of Content, Networks, and Terminal”, 7 mai 2008.

2.6 Séminaires, Tutoriaux et Articles invités

Mars 2004	IRISA –Rennes Séminaire : Transport adaptatif des paquets vidéo sur les réseaux IP NGN
Avril 2004	CRAN – Nancy Séminaire : Marquage et transport des paquets vidéo sur les réseaux IP DiffServ
Janvier 2005	Thales Research and Technology – Reading – UK Tutorial : IP Quality of Service : state of the art and evolution
Septembre 2006	Italy (Macerata) – Lancement du projet ENTHRONE-2 (28 partenaires) Tutorial : Cross-Layer video adaptation

2.7 Thématiques de recherche

Le principal axe de ma thématique de recherche est celui lié au transport des flux vidéo temps réel sur les réseaux de nouvelle génération. L'objectif est d'offrir à l'utilisateur un accès universel à des flux vidéo avec support de la qualité de service, partout, n'importe quand et depuis n'importe quel équipement d'accès. Dans ce sillage, cet objectif recèle des verrous architecturaux sous-jacents qui ont motivé mes travaux de recherche et qui m'ont permis d'apporter une contribution originale à l'édifice.

Dans un contexte de fusion et de convergence des réseaux et des services, l'objectif est de pouvoir mettre en œuvre des services vidéo capables de s'adapter automatiquement aux ressources réseaux disponibles tout en tenant compte des caractéristiques du terminal utilisé, du profil de l'utilisateur et des mécanismes de QoS offerts dans les réseaux sous-jacents. Ainsi, il est nécessaire de caractériser chaque élément pouvant influencer directement ou indirectement la qualité du service perçue par l'utilisateur final et qui est appelée QoE (Quality of Experience). Plusieurs mécanismes de contrôle de QoS doivent être mis en œuvre pour assurer la continuité de QoS à travers différents réseaux de transport et au sein d'une pile protocolaire réseau. Ces mécanismes permettent de maintenir un niveau de QoE acceptable, préalablement défini, et qui peut être évalué par des mesures objectives ou subjectives du flux vidéo reçu. L'adaptation du service doit être réalisée sur plusieurs échelles temporelles et au niveau des différents acteurs de la chaîne du service. Le flux vidéo doit être adapté au début de la session suivant les caractéristiques statiques du terminal et de l'utilisateur puis au cours de la session suivant la variation dynamique des ressources réseaux disponibles et d'autres caractéristiques sous-jacentes. Les acteurs impliqués dans la chaîne de transport des flux doivent harmoniser leur vision pour apporter une solution de QoS de bout-en-bout.

Mes contributions de recherche dans ce domaine peuvent ainsi être classées en trois catégories : (1) le transport et le contrôle de la qualité de service vidéo sur les réseaux IP filaires, (2) le transport et le contrôle de la qualité du service vidéo sur les réseaux IP sans-fil 802.11, et (3) le transport et le contrôle de la qualité de service vidéo sur les réseaux P2P.

Ce domaine de recherche, m'a permis de développer des compétences autour de la gestion et du contrôle de la qualité de service dans les réseaux de transport (IP, sans fil 802.11 et P2P, architecture IETF DiffServ, IntServ), des technologies de l'audiovisuel (codage audio/vidéo MPEGx/H.26x, systèmes de diffusion DVB-T/H, et satellites), ainsi que des architectures de contrôle et de signalisation des services multimédia (3GPP IMS, IETF SIP/NSIS/PBNM, ISO MPEG 21, etc.).

2.7.1 Transport et Contrôle de la qualité des services vidéo sur réseaux IP filaires

Le transport de flux vidéo sur Internet et plus généralement IP est soumis à des contraintes de débit, de délai et de fiabilité. La charge et la nature du trafic sur le réseau ont des effets sur les pertes et les délais d'acheminement des paquets qui doivent être contrôlés. Dans ce contexte, et durant ma thèse de doctorat (2000-2003) dont le sujet était : *Diffusion adaptative des Paquets Vidéo sur IP : une approche Cognitive* », j'ai proposé un système adaptatif de diffusion vidéo (streaming) utilisant des mécanismes de contrôle de QoS appelés « cross-layer ». En effet, l'approche conceptuelle que j'ai proposée supprime la relation hiérarchique entre couches et autorise une meilleure prise en charge de la sémantique et des contraintes applicatives des flux vidéo transportés pour les appliquer dans diverses couches et mécanismes dédiés au transport. Le système proposé est destiné en premier lieu aux réseaux IP fixes qui se conforment au modèle de gestion de QoS par différenciation de services (DiffServ).

Conformément aux modèles architecturaux OSI MPEG-4-part 2 et IETF DiffServ, mes contributions consistent en :

- (i) un protocole de transport temps réel pour flux vidéo MPEG-4 assurant une classification et une mise en paquet automatique des objets audiovisuels, une protection inégale des paquets aux erreurs et un contrôle équitable du débit des sources vidéo (TCP-Friendly);
- (ii) un algorithme de marquage dynamique de paquets vidéo IP pour les routeurs d'accès DiffServ asservi aux variations de bande passante et aux contraintes de qualité de service des flux multimédia transmis.
- (iv) une architecture de passerelle de signalisation permettant à des terminaux IP fixes et mobiles supportant le protocole de contrôle de session IETF SIP (Session Initiation Protocol) d'accéder à des services multimédia MPEG-4 compatible avec la signalisation de contrôle de session OSI MPEG-4 DMIF (Delivery Multimedia Integrated Framework).

Ces contributions ont donné lieu à des publications dans des congrès et revues internationales (IEEE JSAC, Computer Communications Journal, IEEE ICC, IEEE Globecom, IFIP/IEEE MMNS) et à une validation par prototype et expérimentation dans le cadre de projets de recherche financés par la Commission Européenne avec le projet IP ENTHRONE-1 (*Integrated Management of Content, Networks and Terminals*), ainsi que par le ministère de l'enseignement supérieur et de la recherche avec le projet RIAM NMS-*Networked Multimedia Service*.

2.7.2 Transport et Contrôle de la qualité des services vidéo sur les réseaux sans fil 802.11

Cette dernière décennie est marquée par la prolifération des réseaux d'accès sans fil IP, et en particulier les réseaux WLAN de type 802.11 qui occupent une place de plus en plus importante dans les systèmes de communications modernes. Cette prolifération des réseaux sans fil a changé la manière dont les utilisateurs consomment et interagissent avec leurs flux vidéo. Cependant, plusieurs verrous hérités d'IP et d'autres verrous propres aux réseaux sans fil (évanouissement et atténuation du signal, interférences co-canal et bruit, multitude de schémas de modulation et de codage, etc.) doivent être contrôlés avant l'acceptation de ce type de réseaux comme moyens sûrs et fiables de transport de flux vidéo. Ceci nous a motivé à nous pencher sur la problématique du transport de flux vidéo sur les réseaux sans fil 802.11 avec la thèse de Ismail Djama (2005-2008) intitulée « *Adaptation Cross-Layer pour les services vidéo sans fil* ». Vouloir à tout prix assurer une qualité de service a engendré une situation dans laquelle plusieurs mécanismes sont définis au niveau de diverses couches mais sans aucune coordination existante pour assurer la cohérence globale du système. En effet, il existe des mécanismes qui interfèrent les uns avec les autres et qui se dégradent mutuellement, des mécanismes qui s'annulent mutuellement, ainsi que des mécanismes qui ne se propagent pas et restent localisés sur quelques couches protocolaires. Cette situation est doublement frustrante, aussi bien pour l'utilisateur qui consomme le flux, que pour le fournisseur de services.

Dans ce contexte, une nouvelle discipline de recherche est apparue et appelée « la conception cross-layer ». Le modèle cross-layer permet de changer la manière dont les couches protocolaires réseaux échangent leurs données et invoquent leurs services. Ce modèle propose des échanges de messages entre couches pour implémenter un objectif commun tels que : la qualité de service (QoS), la sécurité ou encore la mobilité.

Ceci nous a amené à exploiter ce nouveau concept « Cross Layer » pour faire face aux verrous soulevés en autorisant des échanges d'informations entre les couches. Ce nouveau paradigme suscite un grand intérêt pour améliorer les performances des réseaux sans fil, dont l'état du canal varie considérablement, comparé à un canal filaire. Le partage de l'état du canal physique avec les couches supérieures et applicatives pour prendre en charge efficacement les mécanismes de QoS est un exemple de l'utilisation de la conception cross-layer. Le cross layer permet aussi une collaboration entre les différents mécanismes de QoS qui existent sur les diverses couches protocolaires afin d'assurer une continuité verticale de la QoS pour les flux multimédia.

Les contributions apportées à ce domaine se déclinent ainsi :

- (i) La proposition de mécanismes et de modèles mathématiques permettant l'adaptation du débit vidéo en fonction du débit disponible au niveau de la couche physique 802.11 en assurant un transport fiable par codes correcteurs d'erreurs de type FEC (Forward Error Correction), qui est régi par la variation de la puissance du signal physique et de la mobilité de l'utilisateur.

- (ii) La proposition de mécanismes permettant l'adaptation dynamique de la taille des trames MAC-802.11 pour minimiser les taux de pertes durant la transmission des flux vidéo sur un canal physique comportant un niveau d'interférence important.
- (iii) La proposition de mécanismes permettant aux flux vidéo d'avoir des opportunités de transmission (TXOP) adaptatives en fonction de leur débit grâce à la transmission en rafale (burst) des trames MAC appartenant à la même image vidéo. Ce mécanisme permet ainsi de garantir d'avantage de débit au niveau de la couche physique pour le flux vidéo lorsqu'il est concurrencé par d'autres flux UDP ou TCP, il en résulte une limitation du débit 802.11 due au partage équitable du lien d'accès garanti par le mécanisme d'accès DCF.

Ces travaux de recherche ont été publiés dans des conférences et revues internationales (IEEE Transactions on Multimedia, IEEE Transactions on Broadcasting, IEEE Packet Video), et ont été validés par prototype intégré au démonstrateur du projet Européen IP ENTHRONE-phase 2.

2.7.3 Transport et Contrôle de la qualité des services vidéo sur les réseaux P2P

Dans le paysage de la société de l'information et de la communication, la tendance actuelle repose sur des réseaux et des systèmes d'informations de plus en plus décentralisés, interconnectés, hétérogènes par nature et avec une composante de mobilité et de multimédia très forte. Le passage à l'échelle (scalability), la fiabilité, la performance élevée, la robustesse et la tolérance aux fautes sont des caractéristiques principales à mettre en place pour la survie de ce genre de système. Actuellement, les réseaux P2P sont déployés pour répondre à ces exigences. Les réseaux P2P ont connu un succès rapide, aussi bien auprès du grand public que auprès des professionnels pour mettre en place des systèmes de partage de données simples et efficaces (audio/musique, vidéo, fichier, applications, etc.).

Dans la mesure où cette structuration des réseaux en P2P suscite un grand intérêt et recèle des perspectives prometteuses, nous nous sommes également penchés sur la problématique de la transmission de flux vidéo temps réel sur les réseaux P2P. Un exemple d'utilisation est la diffusion vidéo ou TV sur Internet qui nécessite une infrastructure capable de soutenir un grand nombre de connexions simultanées. Plusieurs verrous restent à lever pour assurer une bonne QoS aux utilisateurs. Ceci nous a motivé pour traiter la problématique du transport de flux vidéo temps réel sur les réseaux P2P avec la thèse de Mubashar Mushtaq (2005-2008) intitulée « *Transport et Adaptation de flux vidéo temps réel sur les réseaux P2P* ».

Nos deux principales contributions sont :

- (i) La proposition d'un mécanisme permettant l'adaptation de QoS du transport de flux vidéo sur les réseaux P2P. Ce mécanisme permet le choix des meilleurs pairs disponibles pour la participation à la session et l'adaptation de la QoS par des techniques de commutation de pairs et de flux.
- (ii) La proposition d'un mécanisme d'organisation des pairs sous forme de réseaux overlay pour optimiser le transport de flux vidéo. Ce mécanisme est couplé à des algorithmes d'ordonnancement et de monitoring de pairs pour assurer un transport lisse (smooth) de flux.

Ces travaux ont été publiés dans des conférences et revues internationales (Journal of Network and Systems Management, IEEE CCNC, IFIP/IEEE MMNS).

2.8 Responsabilités Administratives et Scientifiques

2004 - 2008

- **Coordinateur** des relations internationales pour le département télécommunications à l'ENSEIRB, de septembre 2004 à ce jour.
- **Responsable** de la deuxième année du département télécommunications, année universitaire 2006-2007.
- **Responsable** de la première année du département télécommunications, année universitaire 2005-2006.
- **Membre de la Commission de Spécialistes** (section 27) – Université Bordeaux-1/ENSEIRB/IUT 2007-2008
- **Membre externe du Comité de Sélection**, profil réseaux, Université Paris-V, 2007-2008, recrutement de deux maîtres de conférences sur « profil réseaux ».
- **Co-direction** de thèse avec le Prof. F. KRIEF de quatre doctorants depuis sept. 2005 dont deux d'entre eux ont soutenu leur thèse (U. Javaid et I. Djama)

2.9 Collaborations Nationales et Internationales

- **Université Paris-5 (France)** : équipe du professeur Ahmed MEHAOUA, équipe Informatique et Communication (Infocom), collaboration sur la métrologie des réseaux et la gestion de la QoS des services multimédia sur IP
- **University of Waterloo (Canada)** : équipe du Professeur Raouf BOUTABA : sur la gestion de la QoS des flux vidéo sur IP fixes et mobiles
- **University Politechnic de Bucharest (Roumanie)** : équipe du Professeur Eugen BORCOCI : sur la signalisation dans les réseaux de nouvelle génération
- **University College of Dublin (Irlande)** : équipe du professeur Liam MURPHY, sur le transport P2P des flux vidéo dans les réseaux IP
- **Institut Demokritos - Athens (Grèce)** : équipe du Docteur Anastasios KOURTIS : sur le transport de flux vidéo sur les réseaux DVB terrestres et satellites

- **Thomson Broadcast & Multimédia (Rennes)** : équipe de Marc LE NAOUR sur le traitement et l'adaptation des flux vidéo
- **Centre National d'Etudes Spatiales (Toulouse)** : équipe de Patrick GÈLARD sur la convergence des réseaux et des services dans le contexte des réseaux satellites.

2.10 Financement et Contrats de Recherche

En cours (depuis Sep. 2006) – Fin Prévû en Novembre 2008	Responsable LaBRI du projet européen ENTHRONE2 (6PCRD) “End-to-End QoS Through Integrated Management of Content, Networks and Terminals” . Type du projet : FP6 - IP : Integrated project, http://www.ist-enthrone.org/ Durée : 24 mois du 01 sep. 2007 – 30 Novembre 2008 contrat N° : 38463
Sep. 2008 – Dec. 2008	Responsable LaBRI du projet Bilatérale LaBRI-CNES (Centre National d'Etudes spatiales) - France Titre : “Hybridation du satellite avec les réseaux maillés de capteurs sans fil” Contrat n° : n°4700025329
Nov. 2007 – Mars 2008	Responsable LaBRI du projet Bilatérale LaBRI-CNES (Centre National d'Etudes spatiales) - France Titre : “Les Mécanismes de Gestion de Ressource Cross-Layer” Contrat n° : 71247/00
Nov. 2006 – Mai 2007	Responsable LaBRI du projet Bilatérale LaBRI-CNES (Centre National d'Etudes spatiales) - France Titre : “Integration of Satellite in Next Generation Network for Providing Multicast Multimedia Service” Contrat n° : 60388/00

2.11 Jury de Thèses et Encadrements d'Étudiants

M. Javaid Usman	Date de début: 01/09/2005 , thèse soutenue le 07/03/2008 , mention très honorable Titre : Routing and Mobility Management Strategies in Personal Networking Environments Co directeurs : Toufik Ahmed (40%), Francine Krief (20%), Djamel Meddour (40%) Actuellement “Project Manager” à British-Telecom-UK
M. Djama Ismail	Date de début: 01/10/2005 , thèse en cours, soutenance prévu en Octobre 2008 Titre : Transport et adaptation inter-couches (cross-layer) des services vidéo sans fils Co directeurs : Toufik Ahmed (70%), Francine Krief (30%)
M. Mubashar Mushtaq	Date de début: 01/11/2005 , thèse en cours, soutenance prévu en Novembre 2008 Titre : Diffusion Adaptative des Paquets Vidéo sur les Réseaux P2P Co directeurs : Toufik Ahmed (70%), Francine Krief (30%)
M. Alaeddine Abdallah	Date de début: 02/05/2007 , thèse en cours, Titre : Une solution P2P distribuée pour le transport des flux MultiMedia dans les réseaux Maillés 802.16 Thèse CIFRE à Orange Lab. France Télécom. Co directeurs : Toufik Ahmed (40%), Francine Krief (20%), Djamel Meddour (40%)

- Examineur de Thèse de « Nicolas PALIX » intitulé « Langages dédiés au développement de services de communication » Directeur Charles Consel, soutenu le 17 septembre 2008 à l'université Bordeaux 1
- Tutorat d'étudiants d'Ecoles d'ingénieur 2^{ème} et 3^{ème} année ingénieur, Master et Licence Informatique

2.12 Activités d'Enseignement

Depuis 2004 (MCF) Référence Syllabus ENSEIRB	Intitulé	Nature	
		- C : Cours - EI : Enseignement Intégré - TP : Travaux Pratiques - PR : Projet Avancé	H. eq. TD
RE101	Introduction aux réseaux LAN/WAN	C	18
RE204	Architecture des réseaux	EI	42,5
	Architecture des réseaux	TP	8
RE205	Réseaux et applications répartis	EI	20
	Réseaux et applications répartis	TP	16
RE208	Interconnexion avancée des réseaux	TP	16
RE 332	Réseaux Multimédia et Qualité de Service	C	18
	Réseaux Multimédia et Qualité de Service	TP	16
RE341	La téléphonie : historique, architecture et évolution	C	18
	La téléphonie : historique, architecture et évolution	TP	12
RE342	Réseaux mobiles et sans fil	C	18
PR205	Projet avancé en Télécommunication S4	PR	25
PR307	Projet avancé en Télécommunication S5	PR	35

2.13 Liste des Publications

2.13.1 Dans des Editions de Livres

- [1] George Pavlou, *Toufik Ahmed*, Tasos Dagiuklas, (Ed.s) "Management of Converged Multimedia Networks and Services" proceedings of the IFIP/IEEE International Conference on Management of Multimedia and Mobile Networks and Services (MMNS 2008) edited in **Lecture Notes in Computer Science (LNCS)**, Vol. 5274, ISBN: 978-3-540-87358-7, 171 p, September 2008.

2.13.2 Dans des Chapitres de Livres

- [2] Mubashar Mushtaq and Toufik Ahmed, "Hybrid Overlay Networks Management for Real-Time Multimedia Streaming over P2P Networks" dans **Lecture Notes in Computer Science (LNCS)**, "Management of Multimedia Networks and Services", San José, California, USA, Vol. 4787, pp.1–13, 2007.
- [3] Toufik Ahmed, Ahmed Mehaoua, Raouf Boutaba, Youssef Iraqi, "IP Video Streaming With Fine-Grained TCP-Friendly Rate Adaptation", dans **Lecture Notes in Computer Science (LNCS)**, "Management of Multimedia Networks and Services", Springer-Verlag Editor, vol. 2839, pp. 18-31, 2003.
- [4] Toufik Ahmed, Ahmed Mehaoua et Raouf Boutaba "Dynamic QoS Adaptation using COPS and Network Monitoring Feedback", dans **Lecture Notes in Computer Science (LNCS)**, "Management of Multimedia on the Internet", Springer-Verlag Editor, vol. 2496, pp.250-262, 2002.
- [5] Toufik Ahmed, Guillaume Buridant et Ahmed Mehaoua "Delivering of MPEG-4 Multimedia Content Over Next Generation Internet", dans **Lecture Notes in Computer Science (LNCS)**, "Management of Multimedia on the Internet", Springer-Verlag Editor, vol. 2216, pp.110-127, 2001.
- [6] Toufik Ahmed "Distribution Adaptative de Contenus Audiovisuels sur les Réseaux IP", Chapitre de livre N°6, **Livre Réseaux et Télécoms** (L'autonomie dans les réseaux), édition Hermes, pp. 167-190, septembre 2006.
- [7] Vincent Lecuire, Toufik Ahmed, et Ahmed Mehaoua, "Protection des Flux Audiovisuels MPEG-4 Transportés sur Internet", dans la revue **Technique et Science Informatique: Réseaux et Protocoles**, TSI vol. 23, n° 5-6/2004, pp. 623-650, Ed. Hermes, ISSN 0752-4072, Mai 2004.

2.13.3 Dans des Revues Internationales (avec comité de lecture)

- [8] U. Javaid, T. Rasheed, D. E. Meddour, Toufik Ahmed and N. Prasad, "A Novel Dimension of Cooperation in 4G", in **IEEE Technology and Society Magazine, Special Issue on Potentials and Limits of Cooperation in Wireless Communications: Toward Fourth Generation (4G) Wireless**, Vol. 27, Issue 1, pp. 29-40, Spring 2008.
- [9] Ismail Djama, Toufik Ahmed, A. Nafaa and R. Boutaba, "Meet In the Middle Cross-Layer Adaptation for Audiovisual Content Delivery", in the **IEEE Transactions on Multimedia**, Volume 10, Issue 1, pp. 105 – 120, Jan. 2008.
- [10] Toufik Ahmed and Mubashar Mushtaq "P2P Object-based adaptive Multimedia Streaming (POEMS)" in **Journal of Network and Systems Management**, Special Issue on Peer-to-Peer Technologies in Network and Service Management, Vol. 15, Number 3, pp. 289-310(22), september 2007.
- [11] Borcoci, E. Kormontzas, G. Asgari, A. and Toufik Ahmed "Service Invocation Admission Control Algorithm for Multi-Domain IP Environments", In: Proc. of **Computer Networks: The International Journal of Computer and Telecommunications Networking**. Volume 51, Issue 16. pp. 4669-4678. Elsevier (2007).
- [12] Ismail Djama and Toufik Ahmed "A Cross-Layer Interworking of DVB-T and WLAN for Mobile IPTV Service Delivery" to be published in **IEEE Transaction on Broadcasting**, ISSN 0018-9316, Vol. 53 Issue: 1, Part: 2, pp. 382-390, march 2007.
- [13] Toufik Ahmed, Abolghasem (Hamid) Asgari, Ahmed Mehaoua, Eugen Borcoci, Laure Berti-Équille, and Kormontzas Georgios. "End-to-End Quality of Service Provisioning Through an Integrated Management System for Multimedia Content Delivery" in **Computer Communication Journal SI on Emerging Middleware for Next Generation Networks**, (accepted in 2006), Vol. 30, Issue 3, pp. 638 -651, ISSN:0140-3664, February 2007.

- [14] A. Mehaoua, Toufik Ahmed, H. Asgari, M. Sidibe, A. Nafaa, G. Kormentzas, T. Kourtis, "Service-driven inter-domain QoS Monitoring System for Large-scale IP and DVB Networks", to appear in **Computer Communication Journal, SI on Monitoring and Measurements of IP Networks**, Vol. 29, No. 10, pp. 1687-1695, June 2006.
- [15] Toufik Ahmed, Ahmed Mehaoua, Raouf Boutaba, and Youssef Iraqi "Adaptive Video Streaming Over Next Generation Networks: A Cross-layer Approach" in **IEEE Journal on Selected Areas in Communications (J-SAC)**, Special issue on Intelligent Services and Applications in Next Generation Networks, Vol. 23, NO. 2, pp. 385 - 401 Feb. 2005.
- [16] Toufik Ahmed, Raouf Boutaba and Ahmed Mehaoua, "A Measurement-based Approach for Dynamic QoS Adaptation in DiffServ Networks", in **Computer Communications Journal**, Special issue on End-to-End Quality of Service Differentiation, Elsevier Editor, VOL. 28, Issue 18, pp. 2020-2033, November 2005.
- [17] A. Kourtis, G. Kormentzas, C. Skianis, G. Xilouris, D. Negru, A. Mehaoua, Toufik Ahmed, E. Borcoci, H. Asgari, S. Eccles, E. LeDoeuf "Provisioning of End to End QoS in Diverse Environments: The ENTHRONE View" in **WSEAS Transactions on Communication**, pp 626-631, Issue 2, Volume 3, ISSN 1109-2742, April 2004.

2.13.4 Dans des Conférences Internationales (avec Comité de Lecture)

- [18] Mubashar Mushtaq and Toufik Ahmed, "Smooth Video Delivery for SVC based Media Streaming over P2P Networks" in proc. of IEEE Consumer Communications & Networking Conference (**CCNC 2008**) Special session on P2P, pp. 447-451, Jan 2008.
- [19] Ismail Djama, Toufik Ahmed "MAC-Level Video Frame Grouping using Cross-Layer Architecture" to appear in IEEE International Symposium on Wireless Pervasive Computing (**ISWPC**), pp. 261-265, Santorini Greece, May 2008,
- [20] Mubashar Mushtaq and Toufik Ahmed, "P2P-based Collaborative Media Streaming for Heterogeneous Network Terminals", in proc. of IEEE International Conference on Automation, Quality and Testing, Robotics (**AQTR 2008**), pp. 330-335, May 2008.
- [21] Mubashar Mushtaq and Toufik Ahmed, "P2P-based Mobile IPTV: Challenges and Opportunities" in proc. of the 6th IEEE /ACS International Conference on Computer Systems and Applications (**AICCSA'08**), pp. 975-980, April, 2008.
- [22] Mubashar Mushtaq and Toufik Ahmed, "Multimedia Caching Integration for Efficient Content Delivery over Heterogeneous Networks" to appear in proceedings of the 6th IEEE/ACS International Conference on Computer Systems and Applications (**AICCSA'08**), pp. 1011-1012, April, 2008.
- [23] U. Javaid, D. E. Meddour, T. Rasheed, Toufik Ahmed, "Adaptive Distributed Gateway Discovery in Hybrid Wireless Networks", in the proc. of **IEEE WCNC** conference, pp. 2735 - 2740 Las Vegas, USA, April 2008.
- [24] Ismail Djama, Toufik Ahmed, Daniel Négru, "Adaptive Cross-Layer Fragmentation for QoS-based Wireless IPTV Services", In proc of The 6th **IEEE/ACS International Conference on Computer Systems and Applications**, pp. 993-998, Doha Qatar, avril 2008.
- [25] Mubashar Mushtaq and Toufik Ahmed, "Smooth Video Delivery for SVC based Media Streaming over P2P Networks" in proc. of IEEE Consumer Communications & Networking Conference (**CCNC 2008**) Special session on P2P, pp. 447-451, Jan 2008.
- [26] I. Kofler, C. Timmerer, H. Hellwagner, and Toufik Ahmed, "Towards MPEG-21-based Cross-layer Multimedia Content Adaptation", Proceedings of the 2nd IEEE International Workshop on Semantic Media Adaptation and Personalization (SMAP 2007), pp. 3-8, IEEE Press, London, United Kingdom, pp.3-8 December 2007.

- [27] Toufik Ahmed, Ismail Djama, A.Nafaa, P.Gélard, C.Donny, M.Jérôme, “IMS-based IP Multicast Service Delivery over Satellite Network”, in Procs of The 18th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (**PIMRC'07**), ISBN: 978-1-4244-1144-3, pp. 1-5, Athens Greece, september 2007.
- [28] U. Javaid, D. Meddour, Mahmud, S.A, Toufik Ahmed "Towards Efficient Gateway Discovery in Hybrid Wireless Networks" in Proc. of IEEE Global Information Infrastructure Symposium (**GIIS**), pp.46-51, Marrakech, 2-6 July 2007.
- [29] Ismail Djama and Toufik Ahmed “Adaptive Cross-layer Fragmentation for Reliable Wireless IPTV Services” in proce. of IEEE International Global Information Infrastructure Symposium (**GIIS**), pp. 203-206, Marrakech, 2-6 July 2007.
- [30] U. Javaid, T. Rasheed, D. E. Meddour and Toufik Ahmed, “Personal Network Routing Protocol (PNRP) for Personal Ubiquitous Environments” in IEEE International Conference on Communications (**ICC 2007**), pp. 3100 – 3107, Glasgow, UK, June 2007.
- [31] U. Javaid, T. Rasheed, D. E. Meddour and Toufik Ahmed, “A Profile-based Personal Network Architecture for Personal Ubiquitous Environments” in Proc. of IEEE Vehicular Technology Conference (**VTC-2007 Spring**), pp. 664-668, Dublin, Ireland, April 2007.
- [32] Ismail Djama and Toufik Ahmed “MPEG-21 Cross-Layer QoS Adaptation for Mobile IPTV Service Delivering” in 2nd International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution (**AXMEDIS'06**), 12-15 Dec. 2006. URL : <http://www.axmedis.org/axmedis2006/docs/AXMEDIS2006-programme-ver2-8-FINAL.pdf>
- [33] Daniel Negru, Toufik Ahmed , Ahmed Mehaoua “Ambient Middleware Framework for Efficient Service Provision to Users Anywhere around the Globe” in proc. of Global Telecommunications Conference, **IEEE Globecom'06**, pp. 1-5, 27-30, Nov 2006.
- [34] U. Javaid, D. E. Meddour, T. Rasheed and Toufik Ahmed, “Towards Universal Convergence in Heterogeneous Wireless Networks using Ad Hoc Connectivity” in Proceedings of 9th International Conference on Wireless Personal Multimedia Communications(**WPMC'06**), pp. 464-468, San Diego, September 2006.
- [35] Mubashar Mushtaq and Toufik Ahmed “Adaptive Packet Video Streaming Over P2P Networks Using Active Measurements” in proceedings of the Elevent IEEE Symposium on Computers and Communications (**ISCC'06**), , IEEE Press, Cagliari, Italy. pp. 423 – 428, June 2006.
- [36] Toufik Ahmed, Daniel Negru and Ahmed Mehaoua, “Enabling Adaptive and Interactive Services in Broadcasting Networks”, in **IEEE ICC 2006** Workshop on IP over Broadcasting Networks (IPBN), Istanbul, Turkey, 11 – 15 June 2006. URL: <http://www.ieee-icc.org/2006/index/conference-program-w.html>
- [37] Djamal-Eddine Meddour, Mubasher Mushtaq, Toufik Ahmed, “Open Issues in P2P Multimedia Streaming”, in **IEEE ICC 2006**, workshop on Multimedia Communications Workshop (**MultiCom**), Istanbul, Turkey, 11 – 15 June 2006, URL: <http://www.ieee-icc.org/2006/index/conference-program-w.html>
- [38] Mubashar Mushtaq, Toufik Ahmed, and Djamal-Eddine Meddour “Adaptive Packet Video Streaming Over P2P Networks” in ACM International Conference Proceeding part of International Workshop on Peer to Peer Information Management (**P2PIM**), Vol. 152, Article N°59, Hong Kong, May 2006.
- [39] Toufik Ahmed, Ismail Djama “Delivering Audiovisual Content with MPEG-21-Enabled Cross-Layer QoS Adaptation” in Packet Video 2006 published in **Journal of Zhejiang Univ SCIENCE A** 7(5) from **IEEE Packet Video PV**, pp. 784 784-793. April 2006.
- [40] Ismail Djama, Toufik Ahmed, Daniel Negru, Nawel Zangar “An MPEG-21-enabled Video Adaptation Engine for Universal IPTV Access” in **IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2006**, Las Vegas, NV USA April 6-7, 2006. URL: <http://www.ieee.org/organizations/society/bt/06bmsympo/program.html>

- [41] Daniel Negru, *Toufik Ahmed*, Ahmed Mehaoua "Adaptive Audio-Video Streaming Solution over IP Mobile Environments", in IEEE Wireless Communication Networks Conference (**WCNC'06**), Vol. 4, pp 2156 – 2161, Las Vegas, NV USA, April 4-7, 2006.
- [42] E. Borcoci, A. Asgari, N. Butler, *Toufik Ahmed*, A. Mehaoua, G. Kourmentzas, and S. Eccles "Service Management for End-to-End QoS Multimedia Content Delivery in Heterogeneous" in Advanced Industrial Conference on Telecommunications (**AICT'05**), pp. 46-52, July 2005.
- [43] *Toufik Ahmed*, A. Asgari, E. Borcoci, S. Eccles, G. Kourmentzas, A. Kourtis, A. Mehaoua, G. Xilouris, « Enthroned Core Networking Elements for End-to-End QoS Provision over Heterogeneous Settings » in **IST Mobile & Wireless Communications Summit**, June 2005. URL: <http://www.eurasip.org/Proceedings/Ext/IST05/papers/353.pdf>
- [44] A. Mehaoua, *Toufik Ahmed*, A. H. Asgari, G. Kourmentzas, G. Xilouris "Inter-domain QoS Monitoring System for Supporting Service Assurance and Resource Management across Heterogeneous Networks" in International Workshop on 'Next Generation Networking Middleware' (**NGNM05**), Waterloo, Canada, May 2005. URL : <http://www.cs.uwaterloo.ca/conferences/networking2005/coming/workshop.html>
- [45] C. Skianis, G. Kourmentzas, A. Kourtis, G. Xilouris, A. Mehaoua, *Toufik Ahmed*, O. Negru, "ENTHRONE Perspective for E2E QoS", **IST Mobile & Wireless Communications Summit**, June 2004. URL : <http://www.mobilesummit2004.org/monday.htm>
- [46] *Toufik Ahmed*, A. Mehaoua, H. Asgari, G. Kourmentzas, E. Borcoci, T. Kourtis, and S. Eccles "End-to-End Quality of Service Provisioning Through an Integrated Network Management System", in International Workshop on 'Next Generation Networking Middleware' (**NGNM04**), pp. 37-45, Athens, Greece, May 9-14 2004.
- [47] Abdelhamid Nafaa, *Toufik Ahmed*, and Ahmed Mehaoua "Unequal and interleaved FEC protocol for robust MPEG-4 multicasting over wireless LAN", in IEEE International Conference on Communications (**ICC'04**), Vol. 3, pp. 1431 – 1435, IEEE Press, Paris, June 2004.
- [48] *Toufik Ahmed*, Ahmed Mehaoua and Vincent Lecuire "Streaming MPEG-4 Audiovisual Objects Using TCP-Friendly Rate Control And Unequal Error Protection", IEEE International Conference On Multimedia & Expo (**ICME'03**), Ed. IEEE Press, Volume 2, pp. 317-320, Baltimore, July 2003.
- [49] *Toufik Ahmed*, Abdelhamid Nafaa and Ahmed Mehaoua "An Object-Based MPEG-4 Multimedia Content Classification Model for IP QoS Differentiation" in proceedings of the Eighth IEEE Symposium on Computers and Communications (**ISCC'2003**), Ed. IEEE Press, Volume 2, pp. 1091-1096, Kemer - Antalya, Turkey, June 2003.
- [50] Abdelhamid Nafaa, *Toufik Ahmed*, Yassine Hadjadj Aoul and Ahmed Mehaoua "RTP4Mux: A Novel MPEG-4 RTP Payload for Multicast Video Communications Over Wireless IP" in **IEEE Packet Video Workshop** (PV'03), April 28-29, Nantes, France, April 2003.
- [51] *Toufik Ahmed*, and Ahmed Mehaoua "Adaptive MPEG-4 Streaming Based on Audio-Visual Object Classification And Network Congestion Feedback" in **IEEE Packet Video Workshop** (PV'03), April 28-29, Nantes, France, April 2003.
- [52] *Toufik Ahmed*, Raouf Boutaba and Ahmed Mehaoua "On Enhancing DiffServ Architecture by Dynamic Policy Provisioning Using Network Feedback" **short paper** in IFIP/IEEE Conference on Network Control and Engineering (**Net-Con'02**), Paris, France, pp. 273-278, October 2002.
- [53] *Toufik Ahmed*, Ahmed Mehaoua, Raouf Boutaba, "Interworking Between SIP and MPEG-4 DMIF for Heterogeneous IP Videoconferencing", in IEEE International Conference on Communications (**ICC'02**), Vol. 4, pp. 2469-2473, IEEE Press, New York, April 2002.
- [54] *Toufik Ahmed*, Ahmed Mehaoua and Guillaume Buridant "Implementing MPEG-4 Video On Demand Over IP Differentiated Services" in IEEE International Conference on Global Communications (**Globecom'01**), Ed. IEEE Press, Volume 4, pp. 2489-2493, San Antonio, Texas, USA, November 2001.

- [55] Toufik Ahmed, Guillaume Buridant and Ahmed Mehaoua "Encapsulation and Marking of MPEG-4 Video over IP Differentiated Services" in proceedings of the Sixth IEEE Symposium on Computers and Communications (**ISCC'01**), pp. 346-352, IEEE Press, Hammamet, Tunisia. June 2001.

2.13.5 Dans des Conférences Nationales (avec Comité de Lecture)

- [56] M. Sidibé , A. Mehaoua, Toufik Ahmed et A. Nafaa « Monitoring de Bout-en-bout des Services dans les Futurs Réseaux à Grande Echelle », dans **CFIP'05**, Bordeaux, France, Avril 2005.
- [57] Toufik Ahmed A. Mehaoua, A. Nafaa et R.Boutaba "Gestion Dynamique de la Qualité de Service de Flux Vidéo sur l'Architecture IP DiffServ" dans 5ème Colloque Francophone sur la Gestion de Réseaux et de Services (**GRES'03**), Fortaleza, Brésil, Février 2003.
- [58] Abdelhamid Nafaa, Toufik Ahmed, Yassine Hadjadj aoul, Ahmed Mehaoua "RTP4mux: Un Protocole de Transport et d'Agrégation de Flux Vidéo MPEG-4 pour Réseaux IP Sans Fils" in 5ème Colloque Francophone sur la Gestion de Réseaux et de Services (**GRES'03**), Février 2003.
- [59] Toufik Ahmed, A. Mehaoua et R. Boutaba "Un modèle de contrôle de service de vidéoconférence IP impliquant des terminaux multimédia SIP et MPEG-4" dans **CFIP'02**, Montréal Canada, Mai 2002.
- [60] Toufik Ahmed, G. Buridant, A. Mehaoua et J-P Claude "DVMA: Un Algorithme de Marquage des Flux MPEG-4 sur IP DiffServ" in 4ème Colloque Francophone sur la Gestion de Réseaux et de Services (**GRES'02**), Marrakech, Maroc, décembre 2001.

2.13.6 Autres publications

- [61] M. Mushtaq and Toufik Ahmed "End-to-End Admission Control Mechanism for Multimedia Streaming over SP-driven P2P Networks" in International Conference on Telecommunications and Multimedia (TEMU), on Workshop on Adaptive Multimedia and IPTV Streaming over P2P Networks, July 2008.
- [62] Toufik Ahmed «Diffusion Adaptative des Paquets Vidéo sur IP: Une Approche Cognitive», thèse de doctorat de l'université de Versailles SQY, novembre 2003.
- [63] Toufik Ahmed, A. Mehaoua, R. Boutaba "Interworking Between SIP and MPEG-4 DMIF", **Internet Draft**, IETF, draft-ahmed-dmif-sip-00.txt, work in progress, September 2001.
- [64] M. Sidibé, A. Mehaoua, Toufik Ahmed et A. Nafaa « Monitoring de Bout-en-bout des Services dans les Futurs Réseaux à Grande Echelle », poster à **CFIP'05**, Bordeaux, France, Avril 2005.
- [65] FP6 IST ENTHRONE **Deliverable D21**, co-auteur "ENTHRONE Overall Network Architecture", Mai 2004, disponible sur <http://www.enthrone.org>, Mars 2004.
- [66] Toufik Ahmed (ed.) FP6 IST ENTHRONE **Editor of Deliverable D23**, "End-to-end QoS signaling & policy based management architectures", November 2004, disponible sur <http://www.enthrone.org>, October 2004.
- [67] FP6 IST ENTHRONE-2 **Deliverable D05**, co-auteur, "MPEG-21-based Cross-Layer Adaptation Decision-Taking Engine", disponible sur <http://www.enthrone.org>. March 2007.
- [68] Toufik Ahmed (ed.) FP6 IST ENTHRONE-II **Editor ENTHRONE Deliverable D27**, "Pilot and services integration and tests", disponible sur <http://www.enthrone.org>, February 2007.

3 Transport et Contrôle de la qualité des services vidéo sur IP

3.1 Introduction

La première partie de mes travaux de recherche a porté sur le transport de flux vidéo sur les réseaux IP avec différenciation de services. Ce travail a été mené dans le cadre de ma thèse de doctorat. Le but était de concevoir et d'évaluer un système adaptatif de streaming vidéo sur les réseaux IP DiffServ. Durant ces travaux de recherche j'ai proposé une solution complète de diffusion de flux MPEG-4 sur IP (classification de flux, fragmentation en paquets, marquage sur IP DiffServ, contrôle de congestion, etc.).

Conformément aux modèles architecturaux de la norme MPEG-4 et IP DiffServ, le système développé s'articule autour de trois contributions majeures :

- 1) un modèle de classification automatique d'objets audiovisuels couplé à un algorithme de marquage de flux vidéo appelé DVMA (DiffServ Video Marking Algorithm) [15]¹. Ce modèle basé sur un réseau connexionniste neuronal permet d'associer une priorité relative aux objets audiovisuels MPEG-4 en analysant les méta-données et descripteurs d'objets présents dans les séquences vidéo. Cette assignation automatique de priorité au flux d'objets MPEG-4 facilite, ensuite, le « mapping » de ces flux sur les différentes classes de service DiffServ (PHB) disponibles sur le réseau et leur gestion dans les routeurs en cas de congestion. L'originalité de ce modèle est d'exploiter les indices et les descripteurs d'objets audiovisuels extraits du flux ou disponibles dans les méta-données telles que les coordonnées spatio-temporelles des objets, la cinétique des objets, la taille et la forme des objets, et plus particulièrement les descripteurs de QoS, pour classer intelligemment les objets et leur affecter une priorité relative de bout-en-bout :
- 2) un protocole robuste de fragmentation, d'encapsulation et de transport de flux vidéo avec contrôle d'erreurs et de débits sur RTP, appelé RTP4mux [50]. Ce profil de transport réduit fortement la surcharge de contrôle associée à la couche transport temps réel (RTP), tout en maintenant une synchronisation inter-flux forte et une robustesse face aux erreurs et pertes de données au moyen de codes correcteurs d'erreurs de type FEC. Plusieurs fonctionnalités sont proposées : la synchronisation, la fragmentation, l'encapsulation et le multiplexage optimal des flux élémentaires MPEG-4 composé d'AVO sur les réseaux IP. Ainsi, les objets MPEG-4 exigeant les mêmes critères de QoS, seront regroupés ensemble, encapsuler sur RTP et protégés

contre les erreurs en fonction de leur indice de priorité relative calculé par le modèle de classification précédent (1);

- 3) un algorithme de régulation du débit des sources vidéo réagissant aux congestions sur le réseau et équitable aux autres sources TCP [3]. Cet algorithme permet d'ajuster finement le débit du serveur vidéo par des mécanismes d'ajout ou de suppression d'objets MPEG-4 AVO en fonction de leur priorité relative initialement calculée par le modèle de classification précédent (1).

3.2 État de l'art

La distribution des services audiovisuels et multimédia sur IP est soumise à des contraintes de QoS. La qualité de distribution de ces services n'est toutefois pas satisfaisante de part les différentes contraintes temps réel : bande passante limitée et variable dans le temps, taux de perte élevé, délai de bout-en-bout variable, etc. La charge et la nature du trafic sur le réseau ont des effets sur les pertes et les délais d'acheminement des paquets qui doivent être contrôlés. Dans cette section, nous exposons les principaux problèmes de distribution des paquets audio/vidéo sur IP ainsi que leurs solutions disponibles dans la littérature. Les principaux problèmes cités dans le Tableau 3.1 concernent la bande passante, les pertes et les erreurs de transmission de paquet, la variation du délai de transfert et de la gigue. Evidemment, d'autres problèmes existent mais ils ne seront pas présentés dans cette section.

Problèmes	Solutions proposées
Bande passante	Choix du Codec Techniques d'adaptation
Perte	Gestion des pertes et des erreurs Correction des erreurs Dissimulation des erreurs, etc.
Variation de délai et de la gigue	Mémoire tampon de Playout
Délai de transfert	Gestion de QoS réseau

Tableau 3.1 : Problèmes et solutions liés à la distribution des services audiovisuels sur IP.

3.2.1 Gestion de la bande passante

La bande passante utilisée par un service multimédia peut varier considérablement dans le temps (moment de connexion) et dans l'espace (endroit de connexion). La gestion de la bande passante est liée principalement au choix du codec audiovisuel à utiliser pour un service donné. Ce choix est régi généralement par un compromis entre la taille des unités de données (les échantillons), le délai de compression, le débit à la sortie du codec, et la

¹ La liste des références numériques correspond à l'ensemble de mes publications présentées dans la section 2.13 « Liste des publications »

puissance du traitement (i.e. complexité de l'algorithme). Cette équation à plusieurs inconnues constitue la partie la plus importante dans la conquête du temps réel.

Actuellement, il existe plusieurs normes de codage audiovisuel (voir Figure 3.1). La dernière en date est la norme H.264 développée par le groupe JVT (Joint Video Team), ce dernier regroupait des experts de l'ISO/IEC et de l'ITU-T. Ce codec a été intégré comme partie 10 à la norme MPEG-4. La norme H.264 vise des services très variés (HDTV, VoD, Vidéoconférence, Télévision terrestre et satellite, etc.). Les tests objectifs et subjectifs de la qualité visuelle de H.264 montrent qu'en utilisant deux fois moins de débit par rapport à la norme MPEG-2, H.264 peut atteindre les mêmes qualités visuelles. Cela explique son adoption rapide et large au sein de différents organismes et projets notamment au sein du 3GPP et DVB pour offrir des services audiovisuels pour les réseaux cellulaires de mobiles et les réseaux de diffusions audiovisuelles terrestres et mobiles.

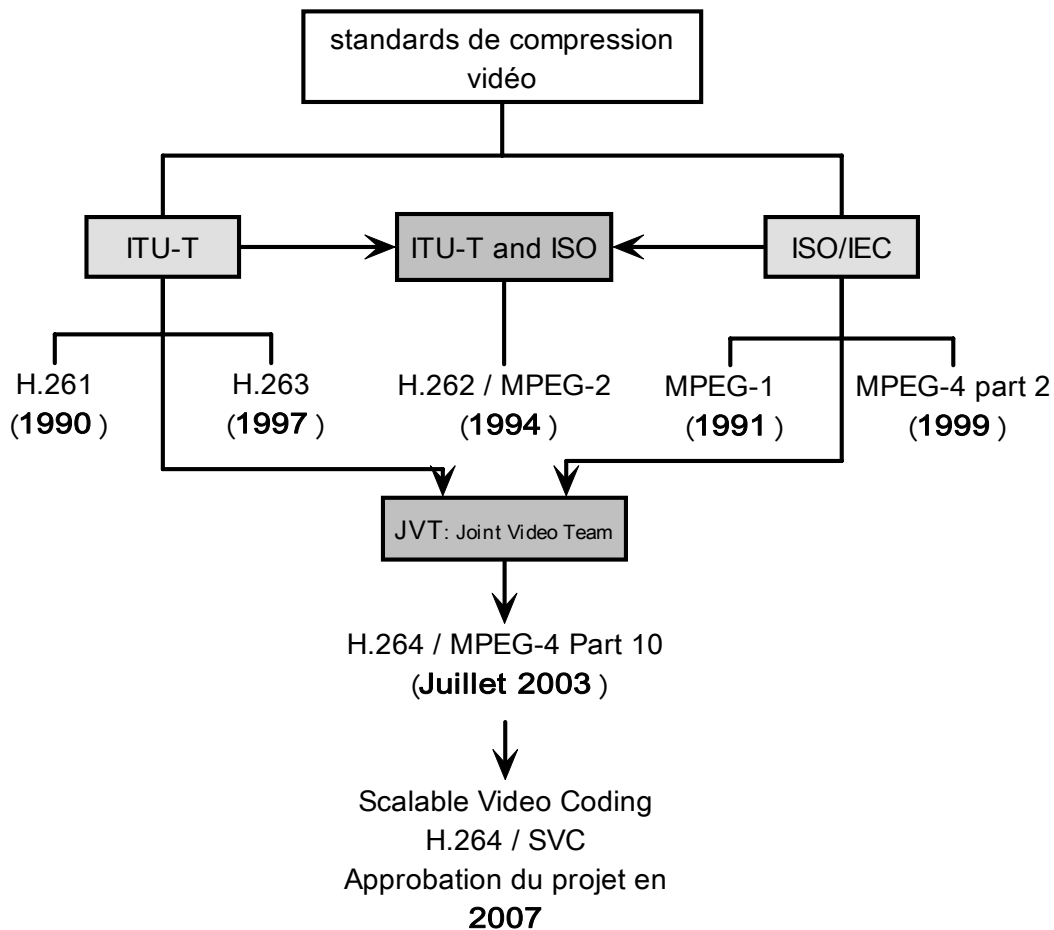


Figure 3.1 : Principales normes de compression audiovisuelles issues de l'ISO et de l'ITU.

Choisir un bon codec répondant au besoin de la bande passante d'un service de diffusion vidéo (exemple : VoD, TV, etc.) ne suffit pas pour déployer un service audiovisuel sur IP. Le débit disponible de bout-en-bout pour le fonctionnement du service peut varier considérablement et à tout moment. Ce phénomène s'aggrave par

L'utilisation des services avec la technique de codage de flux à débit variable (codage VBR). En effet, ce type de codage fournit la meilleure qualité visuelle par rapport à un codage à débit constant (codage CBR), mais requiert plus de contrôle pour assurer une bonne QoS. Les services CBR sont, dans la majorité des cas, non adaptés aux services interactifs et conversationnels (exemple : vidéo conférence, téléphonie) du fait de la stratégie employée qui consiste à utiliser des mémoires tampons pour régulariser et absorber les variations du débit à la sortie du codec. De ce fait, l'utilisation de telles mémoires tampons au niveau de la source augmente le délai de bout-en-bout du service, ce qui peut rendre le service non conversationnel si un certain seuil est dépassé (i.e. si le délai de bout-en-bout dépasse 300ms).

Pour cela, il est nécessaire de gérer la bande passante de manière dynamique en observant l'état des ressources disponibles et en agissant sur les différents éléments impliqués dans le service (c'est-à-dire par adaptation dynamique). Plusieurs techniques de gestion de la bande passante de bout-en-bout permettent à la source multimédia de s'adapter dynamiquement aux changements des conditions réseau. Nous exposons dans ce qui suit quelques unes de ces techniques :

- **SimulStore** (Simultaneous Store and Stream) : cette technique permet de stocker au niveau de la source multimédia (serveur) plusieurs flux audiovisuels. Chaque flux ayant une qualité donnée et fournissant un débit donné. Le client se connecte au serveur de diffusion et négocie avec ce dernier les paramètres de qualité de service de la session. Cette technique est statique et ne permet pas à la source de s'adapter de manière dynamique au changement des conditions réseau.
- **Commutation de flux** (Stream Switching) : cette technique est une extension de la technique SimulStore. Elle procure à la source multimédia des mécanismes permettant de détecter les changements de conditions réseau, d'y remédier en commutant d'un flux à un autre, et ainsi de changer de qualité et de débit sans intervention du client. La synchronisation des flux est assurée par une commutation au niveau des images intra (I) ou d'autres images de type SP (Switching Picture). Une évaluation de cette technique est présentée dans [SUN 02A] [SUN 02B]. Cette technique est aussi employée dans les produits RealVideo développés par real network appelée SureStream [REA 99].
- **SimulCast** (Simultaneous multicast) : dans cette technique la source diffuse plusieurs sessions multicast à plusieurs clients. Chaque session multicast fournit une qualité et un débit donnés. Le client peut estimer les dégradations de sa connexion et peut commuter d'une session multicast à une autre pour s'adapter à ses besoins en termes de bande passante et de qualité.
- **Transcodage** : cette technique peut être employée au niveau de la source audiovisuelle ou au niveau du réseau et peut être déployée sous forme de passerelles appelées « passerelles multimédia » [AMI 95] [AMI 98] et considérées comme des passerelles applicatives (ALG : Application Level Gateway). Le but de cette technique est de changer le débit d'un flux audiovisuel en changeant l'algorithme de

compression utilisé. Par exemple transcoder du format MPEG-2 au format H.264 offre la possibilité de changer le débit en agissant sur la qualité visuelle SNR (rapport signal sur bruit), la résolution spatiale, et la résolution temporelle [JAC 88]. Cette technique nécessite beaucoup de ressources en termes de puissance de calcul et de délai de traitement. Elle n'est pas adaptée aux services directs et interactifs (live) vu les délais supplémentaires qui peuvent être engendrés par les algorithmes de transcodage.

- **Codage conjoint source/canal et utilisation du transrating** : cette technique peut être utilisée pour les services de diffusion directs (live). Elle permet de changer le débit sortant du codec en changeant le facteur de quantification de la matrice DCT (Discrete Cosinus Transform). Grâce cette technique, un module, au niveau de la source audiovisuelle, estime en permanence la qualité du canal de transport (bande passante disponible) et ordonne au module de quantification d'augmenter ou de diminuer la valeur du pas de quantification (Quantizer step). Ceci se traduit par un changement dynamique du débit de la source. Cette technique est appelée « transrating ». Plus le pas de quantification est grand, plus le processus de compression perd des informations au niveau de la matrice DCT, réciproquement, plus le pas de quantification est petit, plus on se rapproche d'un processus de compression sans perte. Augmenter le pas de quantification conduit à un débit faible à la sortie du codec et à une qualité visuelle médiocre, par contre diminuer le pas de quantification conduit à un débit élevé et à une qualité visuelle optimale. La technique du transrating est la plus utilisée de nos jours pour l'adaptation du débit de la source, mais requiert un contrôle approfondi pour estimer correctement l'état du canal et décider du débit de flux.
- **Lissage du trafic vidéo** (video smoothing) : cette technique permet de changer la nature sporadique du trafic vidéo en un flux beaucoup plus stable à débit presque constant pour minimiser les envois en rafale de paquets vidéo. L'envoi de paquets vidéo en rafale est très courant puisque le taux de compression vidéo n'est pas réparti de manière homogène entre les images qui composent un flux. Les images « I » ont besoin de plus de bande passante par rapport aux images « P » qui eux même nécessitent plus de bande passante par rapport aux images « B ». La technique de lissage emploie une mémoire tampon pour réguler le débit à la sortie du codec, mais elle a l'inconvénient de rallonger le délai de transfert dans le réseau, ce qui diminue l'interactivité du service.
- **Streaming audiovisuels avec gestion de bande passante de type TCP-Friendly** : Les applications vidéo utilisent généralement la pile protocolaire RTP/UDP/IP. Or UDP n'applique aucun contrôle de congestion et n'assure pas l'équité du trafic. On peut noter une dégradation importante du débit d'un flux TCP en présence d'un trafic UDP. De plus, le phénomène de congestion est majoritairement causé par les flux UDP. Ce qui présente un problème important pour les flux audiovisuels. Le contrôle de congestion peut être appliqué pour y remédier. Il peut être présent au niveau réseau ou au niveau de la source (politique au niveau routeur, message Source Quench, bit DEC, mécanisme de file d'attente

RED, etc.). Le protocole TCP applique un mécanisme de contrôle de congestion natif au niveau de la source (mécanisme AIMD « Additive Increase and Multiplicative Decrease», slow start / congestion avoidance, fast retransmission / fast recovery, etc.) [JAC 88]. Ainsi, il est nécessaire qu'un trafic UDP applique lui aussi un contrôle de congestion pour être compatible TCP (TCP-Friendly) [MAH 97]. L'idée de TCP-Friendly est introduite pour empêcher une application de rentrer dans une phase où beaucoup de trafic est envoyé et peu correctement reçu. Pour éviter ce phénomène, le trafic qui n'est pas compatible TCP peut être supprimé par les routeurs [FLO 00], ce qui oblige les applications à opérer en mode contrôle de congestion. La montée en charge (scalability) du service audiovisuel est la clé de l'adaptation du trafic à la bande passante. En effet, cette adaptation est basée sur un protocole qui permet d'estimer la bande passante que la source a le droit de consommer. Il existe plusieurs protocoles génériques tels que : LDP « Loss-Delay Based Adaptation Algorithm » [SIS 98], RAP (Rate Adaptation Protocol) [REJ 99], TFRC « TCP-friendly Rate Control Protocol » [HAN 03] qui peuvent être appliqués à cet effet.

Pour résumer, l'étape majeure pour le déploiement d'un service de diffusion multimédia est le choix un bon codec adapté au service et la gestion de la variation possible de la bande passante. Le Tableau 3.2 résume l'ensemble des techniques présentées précédemment et donne une vue globale sur leur domaine d'application et leur degré de complexité.

Techniques	Service Audiovisuel	Complexité
SimulStore	VoD	o
Commutation de flux	VoD	++
SimulCast	VoD	+
Transcodage	Service Live / VoD	+
Codage conjoint source/canal	Service Live / VoD	++
Lissage du trafic vidéo	Service Live / VoD	+
Streaming audiovisuels TCP-Friendly	Service Live /VoD	++

++ : Technique très compliquée

+ : Technique peu compliquée

o : Technique pas compliquée

Tableau 3.2 : Complexité et domaine d'application des techniques de gestion de la bande passante.

3.2.2 Gestion des pertes et des erreurs

Les pertes et les erreurs affectent considérablement la qualité du service perçue par l'utilisateur final. Ils peuvent causer un effet très destructif sur le trafic vidéo à cause des dépendances qui existent entre les images codées (I,

P, et B). Une perte de paquet peut arriver à plusieurs niveaux et dans des contextes différents. Dans les réseaux câblés, la congestion est la première circonstance de pertes de paquets. Un paquet complet peut être supprimé au niveau réseau par un routeur (dépassement des files d'attente, rejet sélectif par les mécanismes de gestion des files d'attentes : RED/RIO, etc.). Le canal de transmission peut causer des erreurs au niveau bit (exemple : interférence), si ces erreurs sont localisées dans les en-têtes du paquet, elles peuvent être détectées par les champs de contrôles CRC et le paquet sera rejeté.

Pour faire face aux effets des pertes sur les flux audiovisuels décodés, il est nécessaire dans un premier temps de détecter et d'isoler la perte, puis dans un second temps d'appliquer des techniques de remèdes. La Figure 3.2 donne un aperçu et une classification des techniques de transport fiable de paquets vidéo.

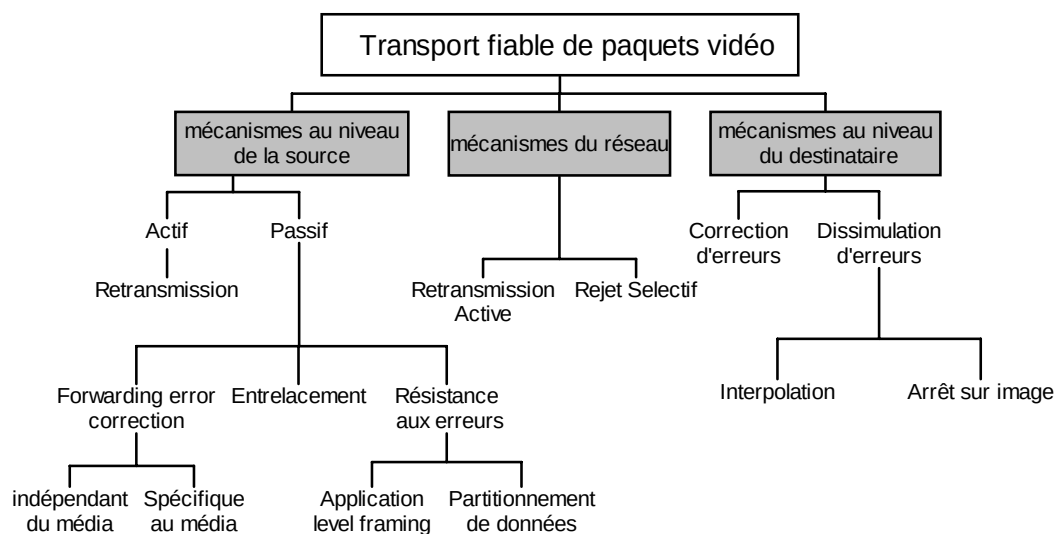


Figure 3.2 : Techniques de contrôle d'erreurs et de pertes.

Ces mécanismes sont classés en trois grandes catégories : mécanismes au niveau de la source, mécanismes au niveau du réseau et mécanismes au niveau du destinataire. Quelques techniques sont détaillées en ce qui suit :

- **Retransmission** : dans cette approche, les paquets perdus sont demandés par le récepteur. Cette technique est appelée dans certain contexte ARQ (Automatic Repeat Request). Pour cela, un canal de retour est nécessaire pour informer la source des paquets à envoyer. L'inconvénient majeur de cette technique est l'augmentation du délai de transfert du paquet perdu à cause des retards introduits par la retransmission. Cette technique s'avère peu utile pour les services directs, en particulier lorsque le délai aller/retour dans le réseau est grand. En cas d'utilisation d'un buffer au niveau du récepteur, une condition peut être introduite consistant à ne transmettre que les paquets pouvant arriver au récepteur avant leur date de consommation. Ceci est rendu possible en analysant l'historique des délais aller-retour dans le réseau ou en utilisant un seuil de tolérance adaptatif qui varie en fonction de la charge du réseau. Plusieurs travaux de recherche exploitent cette technique [MAR 96] [LI 98].

- **FEC (Forwarding Error Correction)** : cette technique est basée sur l'utilisation des codes correcteurs d'erreurs. Des données de redondances sont ajoutées aux flux originaux. Ces dernières sont dérivées des données originales en utilisant des techniques mathématiques basées sur la théorie du codage [LIN 83]. Plusieurs schémas peuvent être appliqués. Le codage Reed Solomon et les bits de parité XOR sont les deux schémas les plus employés. L'inconvénient majeur de cette approche est sans doute l'augmentation du débit de la source à cause de l'overhead introduit par les données de redondance. Le deuxième inconvénient réside dans le retard supplémentaire causé par le stockage des données avant leur traitement par la source pour générer la redondance. Dans le cas des flux audiovisuels, cette technique peut être appliquée en fonction de la charge utile à coder. Elle est dite inégale aux erreurs puisque les parties les plus importantes du flux seront mieux protégées par rapport aux parties les moins importantes [RHE 00] [ADA 03] [LIE 03]. Il est ainsi possible de mieux protéger les images I par rapport aux images P, elles même mieux protégées que les images B.
- **Entrelacement des données** : cette technique vise à réduire l'effet des pertes de paquets en rafale lors d'une congestion sévère dans le réseau. Si une perte a eu lieu sur un paquet donné, il y a une forte probabilité que le paquet suivant soit lui aussi perdu. Il est préférable de disperser ces pertes sur la totalité du flux audiovisuel par rapport à une perte localisée sur un ensemble d'échantillons donné, de telle sorte que les paquets adjacents soient séparés par une distance garantie. L'inconvénient de cette méthode est l'augmentation du délai de latence puisque l'algorithme opère sur plusieurs paquets stockés au préalable avant leur traitement, mais elle conserve le débit du codec puisque aucun paquet de redondance ne sera ajouté. Les deux techniques, à savoir la FEC et l'entrelacement, peuvent être combinées pour une meilleure efficacité face aux pertes de paquets.
- **Résilience aux erreurs** : le partitionnement des données est la technique la plus utilisée dans cette approche. Un paquet vidéo est soit correctement reçu soit entièrement perdu. Ceci signifie que les frontières des données pouvant être perdues sont exactement déterminées par les frontières des paquets. Pour minimiser la propagation des erreurs dans le temps et dans l'espace, il est nécessaire de concevoir un format spécifique de charge utile (payload) pour chaque norme de codage audiovisuel. La source peut mieux organiser son flux en appliquant un bon format de mise en paquet. Par exemple, si le path-MTU du réseau est connu, l'encodeur peut concevoir des paquets ayant la taille d'un MTU et qui soient indépendamment décodable. La plupart des codecs audiovisuels tels que H.263+, MPEG-4, et H.264 supportent cette technique.
- **Dissimulation des erreurs** : afin de diminuer les effets des erreurs sur le flux décodé, il est nécessaire que le récepteur applique des méthodes de dissimulation des erreurs en estimant l'information perdue. La méthode la plus utilisée consiste de cacher le bloc perdu par un bloc équivalent noir ou vert. D'autres techniques plus adaptées aux flux audiovisuels peuvent être appliquées : (1) interpolation (2) arrêt sur

image (3) estimation et compensation de mouvement. L'interpolation est employée pour régénérer un pixel perdu à partir des pixels voisins, c'est l'interpolation spatiale. Dans l'interpolation temporelle le pixel perdu est remplacé par le même pixel se trouvant dans les images suivantes (ou précédentes) correctement reçues. Ces deux techniques peuvent être conjointement appliquées avec une meilleure efficacité si elles sont combinées avec l'estimation et la compensation de mouvement. Ces approches s'appliquent mieux si les pertes sont isolées mais le problème peut s'avérer plus contraignant si plusieurs paquets successifs sont perdus.

Le tableau suivant (Tableau 3.3) résume l'ensemble des techniques présentées précédemment et donne une vue globale sur leurs domaines d'applications et leurs degrés de complexités.

Techniques	Service Audiovisuel	Complexité
Retransmission	VoD	o
FEC	Service Live / VoD	++
Entrelacement des données	VoD	+
Résilience aux erreurs	Service Live / VoD	+
Dissimulation des erreurs	Service Live / Vod	++

++ : Technique très compliquée
+ : Technique peu compliquée
o : Technique pas compliquée

Tableau 3.3 : Complexité et domaine d'application des techniques de gestion des pertes/des erreurs.

3.2.3 Gestion du délai de bout-en-bout et de la gigue

Le délai de bout-en-bout correspond au temps nécessaire pour le transfert de bout-en-bout d'un paquet audiovisuel depuis sa génération jusqu'à sa consommation par le client. La plus grande partie du délai est due au temps de transfert dans le réseau et au buffer de la gigue, mais il inclut d'autres délais à prendre en considération tels que les délais de numérisation, de compression et de mise en paquets. La gestion du délai de bout-en-bout est principalement assurée par les mécanismes de qualité de service réseau. Dans ce contexte, deux grandes architectures existent actuellement. L'architecture IntServ (Integrated Services) accompagnée de son protocole de signalisation RSVP (Resource Reservation Protocol) et l'architecture DiffServ (Differentiated Services). Le développement de DiffServ a été motivé par les besoins des opérateurs pour le déploiement immédiat d'une solution de QoS dans le réseau Internet. Contrairement à la gestion flux par flux de RSVP, les réseaux DiffServ organisent les paquets dans l'une des classes de services identifiées dans un paquet IP dans le champ DSCP (DiffServ CodePoint). Dans chaque routeur DiffServ, un comportement particulier est donné au paquet IP, appelé PHB (Per-Hop Behavior), qui est déterminé par le DSCP. Le comportement total d'un paquet IP reçu par le réseau est considéré comme la somme des comportements individuels de chaque routeur appartenant à

des domaines administratifs distincts. La problématique principale dans l'architecture DiffServ repose sur la gestion et sur la garantie de bout-en-bout d'un comportement donné. En effet, un paquet IP peut traverser plusieurs domaines administratifs et le contrat de service ne peut pas être garanti s'il n'y a pas une vraie collaboration entre domaines. Actuellement, il n'existe pas d'architecture complète de gestion des contrats de service entre des domaines réseaux. Des travaux de recherche sont en cours pour apporter une solution à cette problématique et définir principalement un format de contrat et un protocole de signalisation générique pour négocier et échanger les contrats de services entre les domaines IP DiffServ.

Par ailleurs, la conséquence directe du délai de transfert réseau influence la gigue. Cette dernière concerne le retard engendré entre deux paquets successifs. Elle est significative lorsque le réseau n'est pas homogène. Deux facteurs favorisent l'augmentation de la gigue. Le premier facteur provient du fait que les paquets IP peuvent emprunter des chemins différents pour atteindre une destination donnée. Le deuxième facteur est le taux de remplissage des files d'attente qui engendre un retard dans le traitement de deux paquets successifs. La gigue casse la régularité d'arrivée des paquets au niveau du lecteur du client (player). Si la gigue est importante la mémoire tampon (buffer) au niveau du client se vide rapidement et l'affichage peut s'interrompre à tout moment. Si la gigue est petite, les paquets IP arrivent trop rapidement à la mémoire tampon du client et peut causer un dépassement de capacité. Le dimensionnement de cette mémoire tampon représente un réel challenge pour le déploiement des services audiovisuels. L'utilisation d'une mémoire tampon trop grande rallonge le temps de transfert dans le réseau et rend le service non interactif et non conversationnel si le délai de bout-en-bout dépasse le seuil d'interactivité. D'autre part, l'utilisation d'une mémoire tampon trop petite ne permet pas d'absorber la variation du délai de gigue ce qui augmente les taux de pertes des paquets. La Figure 3.3 illustre le compromis à prendre en considération pour dimensionner une mémoire tampon. Pour un taux de perte de l'ordre de 3%, il faut utiliser une mémoire tampon de 100 ms pour pouvoir accepter les paquets arrivant avec retard de moins de 100ms. Ceci rallonge le délai total du service de 180 ms. Plusieurs travaux de recherche existent pour définir des modèles adaptatifs et analytiques afin de déterminer la taille optimale du buffer de gigue, pour plus de détails voir [FUJ 01A] [FUJ 01B] [STE 01] et [KAL 02].

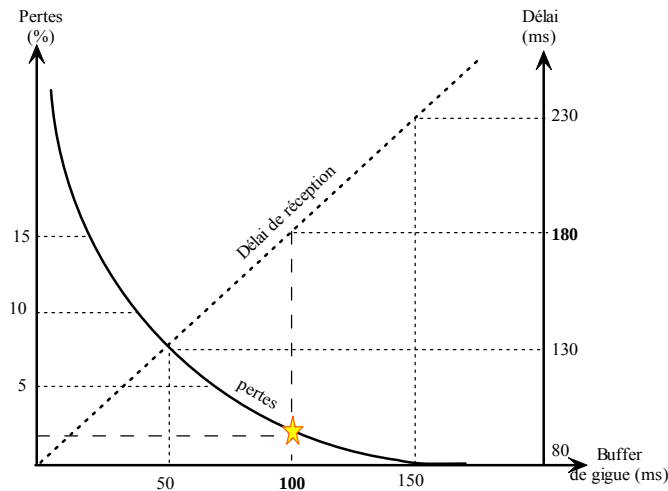


Figure 3.3 : Compromis délai / pertes pour un service multimédia interactif.

L'analyse des différents critères de QoS pour les flux audiovisuels nous a permis de proposer de nouvelles solutions pour le transport et le contrôle de QoS dans un réseau IP. Cette proposition est décrite dans la section 3.3.

3.2.4 Gestion des services audiovisuels : les normes ISO MPEG-4, MPEG-7 et MPEG-21

MPEG-4 : c'est un standard ISO/IEC qui spécifie un ensemble d'outils d'encodage et de traitement de flux audio vidéo naturels et de synthèses. MPEG-4 cible un large spectre d'applications audiovisuelles telles que la vidéoconférence, la vidéo à la demande ou la télédiffusion. L'intérêt principal de ce standard est de fournir un langage commun de description pour l'encodage 2D et 3D des objets audiovisuels conformément aux recommandations des standards : [ISO-IEC 14496.1 98] [ISO-IEC 14496.2 98] [ISO-IEC 14496.3 98].

Une approche orientée objet est adoptée dans le standard MPEG-4 part 2 en vue de représenter des scènes complexes. Une scène est décrite comme une composition d'objets audiovisuels élémentaires, les AVO (*Audio-Visual Objects*), qui sont combinés selon une structure particulière. La Figure 3.4 montre un exemple de scène MPEG-4 part 2 composée de plusieurs AVO, par exemple : un logo, un arrière plan, un présentateur et son discours. L'utilisateur ne perçoit pas de la même manière tous les objets AVO constituant la scène. Dans cet exemple, la qualité du discours audio est plus importante que la qualité d'image du présentateur, elle-même étant plus importante que la qualité de l'image en arrière-plan et du logo. Par conséquent, le transport des AVO sur un réseaux IP peut être différencié pour assurer un niveau de protection en rapport avec l'intérêt que l'utilisateur accorde à chacun d'eux et en concordance avec les mécanismes de transport réseaux implémentés en tenant compte des critères sémantique des flux AVO. Cependant, il faut que la différenciation de service soit

déterminée et connue par les mécanismes de communications sous jacents. C'est l'objet d'une de nos contributions sous la forme d'un modèle de classification automatique d'AVO intégrée à la couche session et qui utilise les méta-données ou descripteurs de flux qui sont supposés être présents dans le flux vidéo MPEG-4.

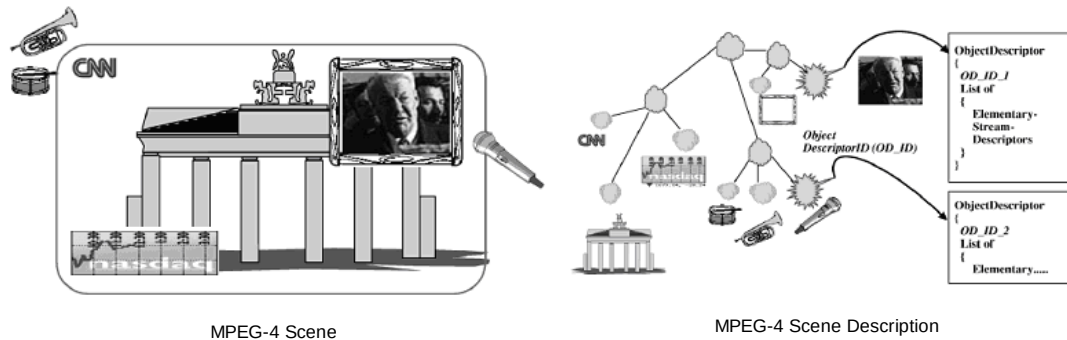


Figure 3.4 : Structure d'une scène MPEG-4 et les descripteurs d'Objets

La couche compression MPEG-4 part 2 génère la représentation des données audiovisuelles dans des flux élémentaires appelée ES (*Elementary Streams*). Les relations hiérarchiques, l'emplacement et les propriétés des ES sont décrits par un ensemble de descripteurs d'objets appelés OD (*Object Descriptors*). Les OD sont eux-mêmes transportés dans un ou plusieurs flux élémentaires. Une scène MPEG-4 est décrite par un descripteur de scène, le SD (*Scene Description*), qui contient les informations sur l'adresse et l'organisation des objets audiovisuels de la scène, en termes d'emplacement spatial et temporel. Le descripteur de scène utilise un langage basé sur le VRML appelé BIFS (*Binary Format for Scene*). Les ES sont fractionnés en une ou plusieurs unités d'accès, les AU (*Access Units*). Une AU représente la plus petite entité de données à laquelle on peut assigner une information de temps. Un flux élémentaire MPEG-4 peut être une vidéo MPEG-1/2 par exemple, et dans ce cas chaque AU va correspondre à une image compressée de type I, B ou P.

Dans la couche Synchronisation, les AU sont transportées sous forme de séquence de paquets appelés SL-PDU. Un SL-PDU est constitué d'un en-tête et d'une charge utile. L'en-tête contient un numéro de séquence pour les besoins liés au contrôle des pertes de paquets et une estampille temporelle pour l'échéance de leur livraison.

Enfin, l'interface entre la couche Synchronisation et la couche de Transfert offre des procédures indépendantes qui permettent d'établir des sessions MPEG-4 et d'accéder à des canaux de transport comme RTP/UDP/IP, AAL/ATM ou DVB.

MPEG-7 : à la différence des autres standards précédemment mis au point par le groupe MPEG (MPEG-1, MPEG-2 et MPEG-4), MPEG-7, ou encore « Multimedia Content Description Interface », ne définit pas la manière dont sont codées les informations. Il normalise un ensemble d'outils de description permettant de définir et d'identifier le contenu de données audiovisuelles : images fixes, vidéo, graphiques, modèles 3D, musiques, combinaisons de plusieurs de ces éléments, etc. Normalisées et intimement liées au contenu, les

descriptions doivent principalement faciliter l'indexation et la recherche dans les environnements multimédia. Cette norme de description, qui devrait être prise en compte par un certain nombre de logiciels et d'équipements de la chaîne multimédia, propose d'aller bien au-delà des systèmes qui permettent actuellement d'effectuer des recherches sur le contenu. Dans la plupart des cas, ces systèmes n'étant pas capables de rechercher une image qu'à partir d'une couleur, d'une forme ou d'une texture. Les outils de description fournis par MPEG-7 sont, en effet, en mesure de préciser un grand nombre d'informations supplémentaires, classées en cinq catégories principales. Elles peuvent bien entendu inclure les informations de bas niveau (couleur, texture, etc.), mais également de type éditorial (titre d'un film, nom du réalisateur, format, etc.), d'éventuelles restrictions sur les conditions d'usage (copyright, etc.), des caractéristiques de stockage et de transmission (format, codage, contraintes de transmission, débits, délais, tolérance aux erreurs/pertes, etc.), de structures spatiales, temporelles ou spatio-temporelles (zones de déplacement d'un objet, etc.).

MPEG-7 a été développé, notamment, en tenant compte d'autres projets plus spécifiques, tels que SMPTE, TV-Anytime, EBU P/Meta et Dublin Core. Le choix d'un langage pour la description des contenus s'est porté sur le modèle XML Schema, normalisé par le W3C. Dans cette optique, MPEG-7 standardise trois types d'éléments principaux. Les descripteurs (D) qui définissent la syntaxe et la sémantique à utiliser pour représenter les caractéristiques propres au contenu. Ils peuvent être de type visuel pour tout ce qui concerne les caractéristiques de couleur (dominante, histogramme, etc.), de texture, de forme ou de mouvement (de la caméra, d'un objet, d'une trajectoire, etc.). Ils peuvent être également de type audio pour définir, par exemple, le timbre d'un instrument, la voix d'un orateur (associée à des méta-données) ou le rythme et le profil d'une mélodie. Le deuxième type d'éléments standardisés par MPEG-7 sont les « *Description Schemes* » (DS). Ils prédéfinissent la structure et le cadre des relations qui seront établies entre différents composants (descripteurs, ou Description Schemes). Enfin, le « *Description Definition Language* » (DDL), permet de créer ou de modifier les Description Schemes.

Ainsi l'originalité de notre activité de recherche dans la conception d'un cadre de transmission de flux vidéo sur les réseaux IP, a été de coupler les descripteurs de QoS de niveau applicatif (descripteur d'objet MPEG-4, MPEG-7) et les différents profils (utilisateur, terminal, réseaux d'accès) avec les algorithmes de contrôle de la qualité de service présents au niveau des couches Transport et Réseaux.

MPEG-21 : Au début des années 2000, le groupe MPEG a décidé de se lancer dans la gestion de la chaîne de service audiovisuel. Actuellement, plusieurs éléments existent pour établir une infrastructure de distribution et de consommation de contenus multimédia. Cependant, il n'existe aucune vue d'ensemble pour décrire comment ces éléments existants ou en cours de développement, se relient et s'ingèrent entre eux pour assurer la meilleure qualité de service à l'utilisateur final. Cette problématique est en cours de définition et de standardisation au sein du groupe de travail MPEG-21. Cette future norme décrit comment ces divers éléments s'adaptent ensemble. Le standard MPEG-21 devrait proposer un cadre ouvert pour la distribution et la consommation de contenu

multimédia agissant sur l'ensemble de la chaîne du service, depuis la génération du contenu jusqu'à sa consommation par les clients.

Sept éléments fondamentaux ont été identifiés dans l'architecture MPEG-21, voir Figure 3.5. Ces éléments fonctionnent ensemble pour fournir un environnement sécurisé et contrôlable, capable de fournir un contenu multimédia riche aux utilisateurs finaux (un flux multimédia est appelé digital DI « Digital Item » dans la terminologie MPEG-21), dans un format approprié aux dispositifs des utilisateurs, tout en satisfaisant leurs besoins en terme de QoS, d'interopérabilité et de sécurité.

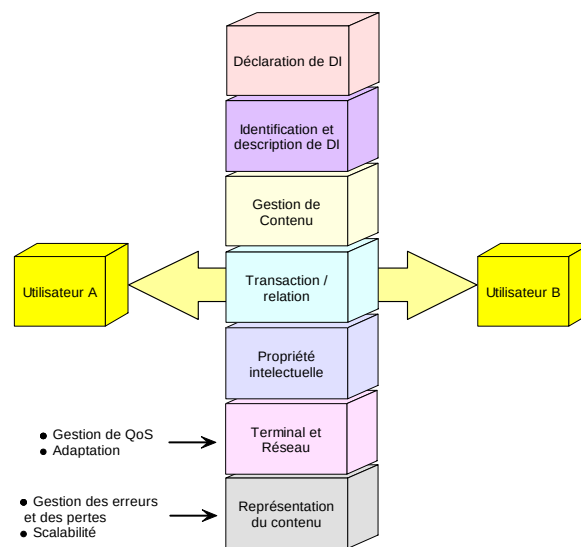


Figure 3.5 : Les éléments architecturaux de MPEG-21.

L'élément clé de l'architecture MPEG-21 est la possibilité de distribution de contenu multimédia pour une multitude de terminaux et à travers plusieurs types de réseaux. Ce concept d'accès universel au média (**UMA** : Universal Multimedia Access) est rendu possible par l'utilisation d'une machine spécifique d'adaptation de contenu multimédia appelé DIA (Digital Item Adaptation). Le concept d'adaptation est sujet à deux processus : (1) processus d'adaptation de ressource qui permet de rendre les ressources utilisables sur n'importe quel type de terminal et de réseau, et (2) processus d'adaptation de descripteur des ressources qui décrivent le plus précisément possible une ressource adaptée. La Figure 3.6 illustre comment adapter un flux élémentaire en utilisant la machine d'adaptation définie dans MPEG-21.

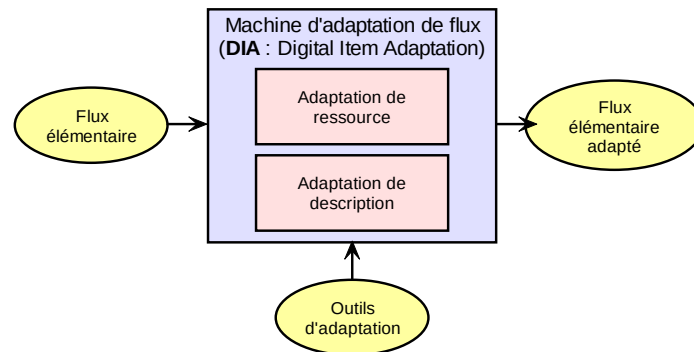


Figure 3.6 : Machine d'adaptation dans MPEG-21.

Nous avons exploité les outils MPEG-21, à savoir les descripteurs de l'environnement de l'utilisateur appelé UED (Usage Environment Description), en particulier les profils du terminal et de l'utilisateur pour adapter le flux vidéo au début de la session par l'implémentation d'une machine d'adaptation de flux.

3.2.5 Les profils de transport temps réel (RTP) pour les services vidéo MPEG4

Plusieurs propositions de profil de transport (RTP) pour des flux MPEG-4 sur réseaux IP ont été soumises au groupe de standardisation IETF AVT (Audio Visual Transport WG), voir Figure 3.7. La première à avoir été retenue, et qui est devenue depuis un RFC sous le n° 3016 [YOS 00], préconise des mécanismes de fragmentation des flux audio/vidéo MPEG-4 pour ensuite les encapsuler dans un flux RTP sans utiliser le système MPEG-4 SL (Sync Layer) qui, à l'instar de RTP, offre des services de synchronisation des flux élémentaires (ES) MPEG-4 indépendamment des caractéristiques du réseau sous-jacent. Le RFC 3016 fournit des règles pour l'utilisation des champs de l'en-tête RTP et la fragmentation du flux vidéo en entités indépendamment décodables. Il donne enfin des indications pour l'utilisation des types MIME (Multipurpose Internet Mail Extension) et de SDP (Session Description Protocol) [HAN 98] pour la signalisation et la configuration des types et paramètres des médias utilisés dans la session MPEG-4 au niveau du terminal de réception. L'ensemble des approches soumises au groupe AVT sont illustrées dans la Figure 3.7.

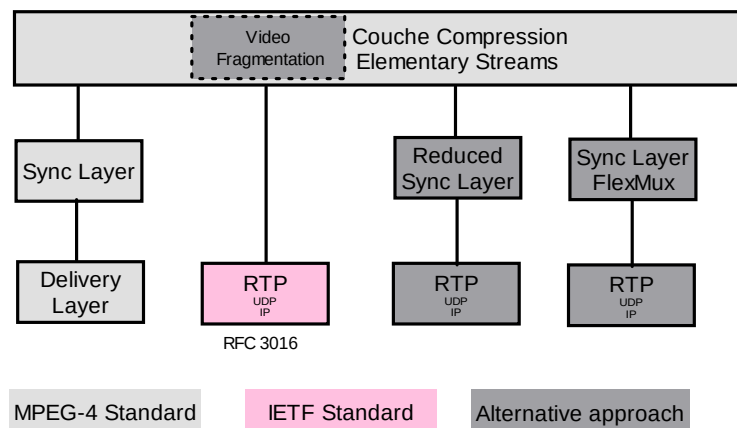


Figure 3.7 : Transport des flux MPEG-4 avec un en-tête SL réduit

Dans l'approche RFC 3016, les flux MPEG-4 (audio et vidéo seulement) sont directement encapsulés dans des paquets RTP, après fragmentation préalable. Chaque flux élémentaire est acheminé dans une session RTP individuelle, identifiée par une adresse réseau et un numéro de port. Le RFC 3016 ne spécifie pas l'utilisation des flux élémentaires du système MPEG-4 (OD, BIFS, ..., etc.), mais ceci peut être réalisé par les applications à travers, par exemple, le protocole de signalisation H.245. Ceci représente un atout pour le déploiement d'un service de streaming impliquant des flux MPEG-4 au côté d'autre média non MPEG-4.

Une seconde approche, proposée à l'IETF et décrite dans [BAS 02] et [VAN 04], préconise en revanche l'utilisation de la couche SL, puis l'encapsulation du flux de paquets SL dans le payload RTP avec un éventuel mapping, de quelques champs redondants, de l'en-tête SL (voir Figure 3.7) vers l'en-tête RTP. L'approche peut être implémentée en utilisant le temps d'envoi comme étant le temps de composition venant de l'horloge de référence dans le champ RTP timestamp. Ce mode d'encapsulation fait intervenir un mapping (transfert) de certains champs pertinents de l'en-tête SL vers l'en-tête RTP. Parmi les avantages de cette approche, nous pouvons citer : (1) un payload générique qui ne dépend pas de la nature du flux transporté (audio, vidéo, OD, SD, ..., etc.), (2) un payload hautement configurable (i.e. la structure du payload RTP dépend des paramètres MIME) et ajustable pour convenir à chaque flux élémentaire individuel, (3) deux styles d'encapsulation (singulière et multiple). L'une permet la compatibilité avec le RFC 3016, tandis que l'autre exploite mieux la bande passante en diminuant l'overhead par l'encapsulation (concaténation) de plusieurs AU dans un même payload RTP, (4) la possibilité d'entrelacer puis de transporter plusieurs fragments d'AU dans un même payload RTP, ce qui permet une plus grande tolérance aux pertes de paquets RTP.

Enfin, une troisième approche, également proposée à la standardisation par l'IETF et décrite dans [CUR 01], repose sur l'encapsulation d'un nombre entier de paquets FlexMux (outil défini par MPEG-4 pour le multiplexage des flux) complets dans la charge utile RTP. Contrairement aux autres approches, proposées par l'IETF ([YOS 00], [BAS 02] et [VAN 04]) qui s'appuient sur l'utilisation d'autant de sessions RTP qu'il existe de

flux élémentaires dans la session MPEG-4. Cette approche permet le multiplexage de plusieurs ES sur une même session RTP, ce qui entraîne une utilisation de moins de ressources systèmes au niveau des serveurs (nombre réduit de sockets ouvertes). Les autres avantages sont : (i) l'utilisation d'une session commune pour des flux qui sont étroitement synchronisés à travers la table codetable du FlexMux (table qui doit être signalée au récepteur au préalable) ; (ii) l'utilisation de la technologie FlexMux permettant une interconnexion entre le réseau Internet et les réseaux de la télévision numérique DVB (avec l'utilisation de la syntaxe FlexMux de MPEG-4 pour transporter des flux MPEG-4 sur les canaux de transport MPEG-2) ; (iii) possibilité de regrouper des flux élémentaires (ES) MPEG-4 ayant les mêmes niveaux d'exigences en QoS, dans un même flux FlexMux, puis dans une unique session RTP, pour réaliser un mapping efficace de la QoS sur le réseau de transport sous-jacent (comme par exemple l'agrégation de plusieurs flux vidéo MPEG4 sur un PHB DiffServ).

Ainsi, nous pouvons distinguer 3 possibilités pour le multiplexage des ES dans les flux RTP :

- ***Une session RTP pour chaque ES:*** ce mode présente l'avantage de pouvoir bénéficier de la possibilité de différencier les flux élémentaires selon leurs besoins en QoS. De plus, les ES_IDs sont signalés à l'établissement de la session MPEG-4 et ne nécessitent pas une autre signalisation ou une maintenance d'état au cours de la session. Cependant, vu que les ES sont dynamiques dans une session MPEG-4 part 2, l'apparition d'un nouveau flux élémentaire dans la session MPEG-4, introduit des délais supplémentaires pour l'établissement de la session RTP associées à un ES.
- ***Une session RTP pour tous les ES:*** ce mode offre une gestion simplifiée des ES mais ne permet pas la différenciation entre les flux élémentaires selon leurs critères en QoS. Il est nécessaire au marqueur de QoS au niveau IP (exemple : DiffServ IP), de remonter jusqu'au niveau système et applicatif pour pouvoir différencier les éléments ES entre eux. C'est le cas du RFC 3016.
- ***Une session RTP pour quelques ES (Agragation des ES):*** ce mode permet, en plus de la gestion simplifiée des ES, l'agrégation et le regroupement des ES sur la base d'une classification préalable et sur des considérations de QoS, en permettant une différenciation des différents regroupements et l'exploitation du codage hiérarchique. Ce mécanisme a été retenu pour notre schéma de transport appelé RTP4mux qui sera présenté dans la section 3.3.

L'analyse et la comparaison de ces différentes propositions de transport RTP nous ont permis de proposer un profil de transport RTP optimal pour des flux audio et vidéo MPEG-4. Cette contribution est appelée RTP4mux et est décrite dans les sections suivantes.

3.3 Contributions

3.3.1 Un modèle de Classification automatique d'Objets Audiovisuels MPEG-4

La classification des AVO est assurée par un modèle à deux niveaux. Le premier niveau correspondant à une classification des AVO selon leurs importances dans la scène, et le second correspond à une classification inter-AVO, c'est-à-dire entre les unités qui composent le flux d'un même AVO.

Dans l'architecture MPEG-4, un AVO est décrit par un descripteur d'objet (OD) qui spécifie la configuration du flux de données. Dans un OD, plusieurs descripteurs coexistent, on y trouve par exemple *DecoderConfigDescriptor* qui décrit le type du décodeur utilisé ainsi que les ressources nécessaires pour bien interpréter le flux de données au niveau client. Le descripteur qui nous intéresse pour les besoins de la classification des flux est celui qui décrit les paramètres de QoS demandés par chaque objet MPEG-4. Ce descripteur est appelé *QoSDescriptor*. Le « *QoSDescriptor* » encapsule des paramètres de qualité de service nécessaire pour un objet AVO. Le Tableau 3.4 présente les principaux descripteurs de QoS associé aux objets MPEG-4.

Nous exploitons les informations contenues dans ce descripteur pour classer les différents objets existants dans une scène MPEG-4. Nous utilisons pour cela un algorithme de classification à base de réseaux de neurones asservi par les descripteurs de classe de service DiffServ.

La technique de codage utilisée dans MPEG-4 permet le codage de chaque AVO par un codeur adapté pour obtenir le meilleur ratio de compression. Dans une scène MPEG-4 part 2, on peut donc trouver un objet vidéo codé en H263 et un autre en MPEG-1 par exemple.

Métrique	Description
MAX_DELAY	Délai maximum accepté de bout-en-bout
PREF_MAX_DELAY	Délai maximum préférable de bout-en-bout
LOSS_PROB	Probabilité de perte des AU acceptée
MAX_GAP_LOSS	Nombre de perte d'AU consécutives acceptée
MAX_AU_SIZE	Taille maximale d'une AU
AVG_AU_SIZE	Taille moyenne d'une AU
MAX_AU_RATE	Débit maximal des AU
PRIORITY	Priorité du flux

Tableau 3.4 : Les descripteurs MPEG-7 de qualité de service associés aux AVO

Le codage MPEG génère trois types d'images codées, fournissant trois niveaux de « scalabilité » temporelle qui sont : I-VOP (*Intra coded Video Object Plane*), P-VOP (*temporally Predicted Video Object Plane*), et B-VOP (*Birectionally predicted Video Object Plane*).

Le codeur MPEG autorise la configuration de la fréquence et de la position des images clés. Cette configuration va déterminer comment les images sont rassemblées en groupes appelés GOP (*Group Of Pictures*). Une séquence

vidéo étant considérée comme une suite de GOP. Un GOP est constitué de N images. Il commence par une image de type I suivie par des séquences constituées par un nombre $(M-1)$ d'images B et d'une image P. Par exemple, une configuration fixée à $M=3$ et $N=12$ donne des groupes d'images formés sur la série IBBPBBPBBPBB. L'importance des AU appartenant aux différents types d'images n'est pas la même. L'image de type I-VOP est la plus importante puisque sans elle, c'est tout le GOP qui ne pourra pas être reconstitué correctement. Les images de type P-VOP ont une importance qui dépend de leurs positions dans le GOP.

D'une manière générale, les sources vidéo MPEG-4 sont codées selon un modèle de représentation multi niveaux. La source vidéo est codée en plusieurs flux compressés correspondant chacun à un incrément de débit et de qualité en termes de corrélation temporelle, de résolution spatiale, ou de rapport signal sur bruit. On distingue un flux de base (BL ou Base Layer), qui est le plus important, et un ou plusieurs flux d'amélioration (EL ou Enhancement Layer). La norme H.264 SVC (Scalable Video Coding) standardise ce modèle de représentation de flux hiérarchiques pour cibler des réseaux et des terminaux hétérogènes. La Figure 3.8 illustre la présentation des couches (appelées aussi des tiers) dans le standard SVC. Plusieurs niveaux de qualité peuvent être transmis à différents réseaux et terminaux. La couche de base, la plus importante, permet de reproduire un flux avec une qualité QCIF et 7.5 images par seconde. L'ajout des couches augmente la qualité du flux sur trois axes : Qualité temporelle (7.5, 15, 30), qualité spatiale (QCIF, CIF, 5CIF), et qualité SNR.

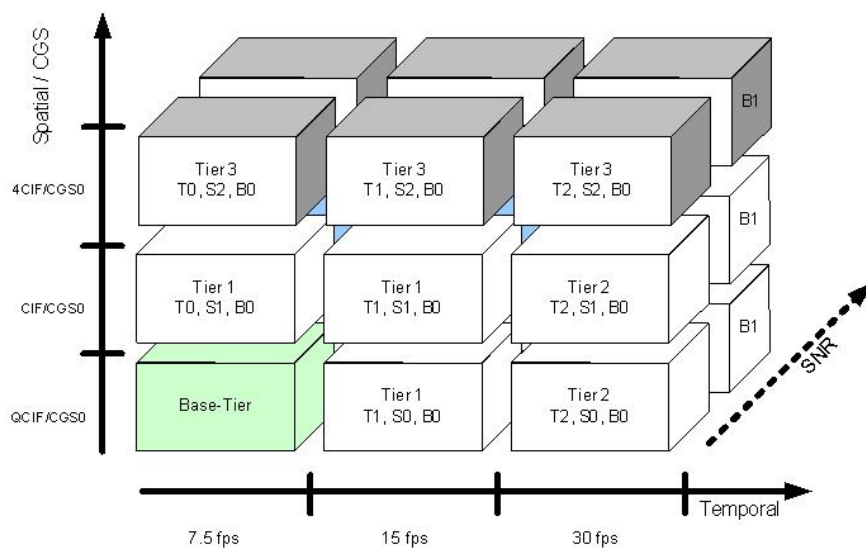


Figure 3.8 : Représentation des couches dans H.264 SVC (Scalable Video Coding)

Nous avons utilisé un algorithme de classification basé sur la technique de classification RBF (*Radial Basis Function*). RBF permet de trier un ensemble d'objets en utilisant une fonction basée sur un vecteur de distance. Nous exploitons la caractéristique du réseau de transport sous-jacent (DiffServ dans notre cas) pour marquer chaque objet avec une étiquette (un label), qui définit la classe de service de l'objet.

Considérons un objet MPEG-4 à classifier, caractérisé par un vecteur $\mathbf{X} \in \mathfrak{R}^n$. Ce vecteur $\mathbf{X}$ est défini par $(x_1, x_2, \dots, x_n)^T$, où chaque composante x_i représente une caractéristique de qualité de service associée à l'objet, parmi celles données par le Tableau 3.4 ou celles disponibles dans les méta-données. Et T est l'ensemble des étiquettes qu'un objet peut recevoir. Ceci correspond aux différentes classes de service DiffServ définies dans le réseau ou à un degré de priorité. Notre système de classification correspond à une fonction $C : \mathfrak{R}^n \rightarrow T$ qui assigne une étiquette (ou une priorité) $c \in T$ à chaque vecteur $\mathbf{X} \in \mathfrak{R}^n$ en se basant sur un ensemble de vecteurs caractérisant la modélisation des classes de service DiffServ ou des priorités données par les modèles $\{U_1, U_2, \dots, U_m\}$.

La Figure 3.9 illustre le principe de fonctionnement du classificateur RBF. La méthode utilise un ensemble de blocs reliés entre eux pour former un réseau. Le chemin parcouru de l'entrée à la sortie du réseau est appelé neurone. Un bloc va correspondre à un traitement particulier identifié par une fonction de transfert. Cette fonction doit prendre en compte les caractéristiques de l'objet MPEG-4 en entrée.

La fonction de transfert $\varphi_{j=1,2,\dots,n}$ est une fonction gaussienne centrée sur un point de l'espace d'entrée. Pour une entrée donnée, chaque sortie du bloc φ_j est calculée selon l'équation suivante : $\varphi_j(x) = e^{-\frac{D^2}{2\sigma_j^2}}$.

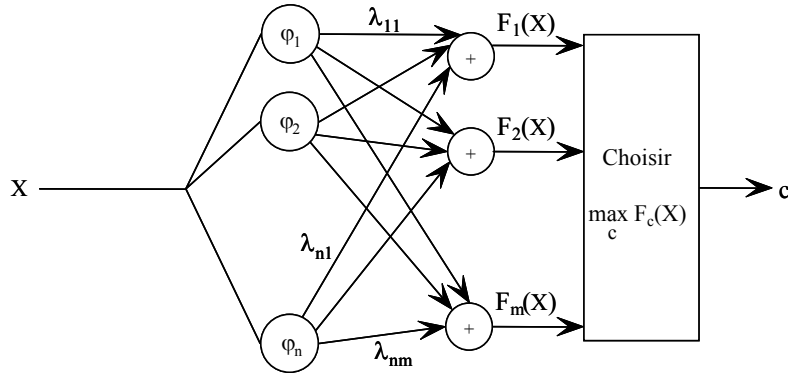


Figure 3.9 : Architecture du classificateur RBF

Le terme $D = \|\mathbf{X} - U_k\|^2$ représente la distance entre le vecteur $\mathbf{X}$ et le vecteur modèle U_k pour la caractéristique de rang j et σ_j est la déviation standard permettant de régler la sensibilité du résultat aux données d'entrées. Chaque étape de processus de classification emploie une donnée pour modifier les poids de toutes les connexions. La capacité d'apprentissage de ces systèmes est liée à la modification des poids des connexions à partir d'exemples, ce qui permet à notre système d'être réactif face aux erreurs de transmission de bout-en-bout. Ce modèle peut être couplé à des mécanismes de « **monitoring de QoS** » pour permettre un apprentissage de la qualité fournie à l'utilisateur final asservi par la variation de la QoS au niveau réseau.

La fonction gaussienne permet aux neurones de ne répondre qu'à une petite région de l'espace d'entrée, région sur laquelle la gaussienne est centrée. La sortie du réseau est simplement une combinaison linéaire des sorties des neurones RBF multipliés par le poids λ_{jk} de leur connexion respective. Le poids permet à l'utilisateur de donner plus d'importance à certains paramètres de qualité de service par rapport à d'autres (par exemple favoriser le délai par rapport au taux de perte).

Dans le classificateur RBF, l'équation générale du résultat d'un neurone est donnée par la relation :

$$F_k(X) = \sum_{j=1}^n \lambda_{jk} \phi_j(X). \text{ Le classificateur choisit l'étiquette correspondant au résultat du neurone le plus grand.}$$

La connaissance des caractéristiques du réseau de transport sous-jacent (i.e. les classes de service) est primordiale pour la classification des objets AVO. Le monitoring des classes de service permet d'assister le système vers la fourniture d'une meilleure qualité [15].

3.3.2 RTP4mux : Un profil de transport RTP pour le multiplexage de flux MPEG-4

Un protocole d'encapsulation efficace réduisant l'overhead (surcharge d'en-tête de contrôle protocolaire) est essentiel pour le transport optimal des flux MPEG-4 sur les réseaux IP. Ce protocole doit minimiser le problème de surcharge due aux différents en-têtes d'encapsulation. Le Tableau 3.5 illustre ce problème et donne quelques exemples d'overhead pour deux flux audio (AAC et CELP). Le protocole RTP4Mux spécifie un format de charge utile RTP générique intégrant des fonctionnalités de gestion temporelle existantes dans la couche MPEG-4 Sync Layer (Temps de composition « CTS » et temps de décodage « DTS »). Le format de charge utile doit, également, permettre le multiplexage de plusieurs flux élémentaires dans une même session RTP pour faciliter la gestion des flux élémentaires, exploiter convenablement la charge utile de RTP et favoriser la mise à l'échelle (scalability) du côté serveur.

Audio Object	Audio Payload Length	Overhead (RTP/UDP/IP)	Overhead Rate
AAC (64 kbit/s, 24 kHz)	342 bytes (average)	12+8+20 bytes	10,9%
CELP (6 kbit/s, 8 kHz)	15 bytes (fixed)	12+8+20 bytes	72,7%

Tableau 3.5 : Exemple d'overhead introduit par le transport de flux MPEG-4 à bas débit

Lorsque le codage AAC (Adaptive Audio Coding) est utilisé pour un signal sonore stéréo à 64 Kbps, la couche de compression génère des trames AAC avec une taille moyenne de 200 octets. Pour optimiser l'utilisation de la bande passante et diminuer l'overhead, plusieurs AU doivent être concaténées ensemble, en particulier pour les flux audio ayant des AU de petite taille. Par exemple pour un réseau WLAN ayant 1500 octets de MTU, 7 trames AAC peuvent être concaténées ensemble, ce qui fournit un gain énorme en bande passante. Cependant, le délai

de bout-en-bout augmente en fonction du nombre d'unités à concaténer, car il faut attendre la génération des 7 trames pour les envoyer sur un seul paquet de transport.

Dans les flux à bas débit (AAC, FA « Facial Animation », CELP « Coding Expedited Linear Predictive audio », TTSI « Text-to-Speech Interface », etc.), l'intervalle temporel entre la génération de 2 AU successives peut s'avérer très important. La concaténation de plusieurs AU de ce flux, pour diminuer l'overhead, impliquera des délais de mise en paquets très importants et intolérables pour les services conversationnels interactifs dont le délai de bout-en-bout ne doit pas dépasser 200 à 300ms. Les flux MPEG-4 bas débit sont destinés à être très largement déployés dans les réseaux Internet et WLAN, d'où la nécessité d'optimiser la bande passante à travers une bonne exploitation du payload RTP.

À partir de ce constat, nous avons proposé l'encapsulation, dans le payload RTP, de plusieurs AU appartenant à différents flux élémentaires. Nous allons voir, par la suite, que nous regroupons les flux ayant les mêmes exigences de QoS dans un même paquet de transport RTP. Ceci permet d'augmenter l'efficacité de l'encapsulation par la réduction du délai de groupement de plusieurs AU et, par la même, l'optimisation de la QoS. Nous pouvons, ainsi, palier au problème de la faible exploitation de la charge utile de RTP et le gâchis de la bande passante entraîné. La Figure 3.10 illustre le processus d'encapsulation des flux MPEG-4 ES proposé dans l'approche RTP4mux.

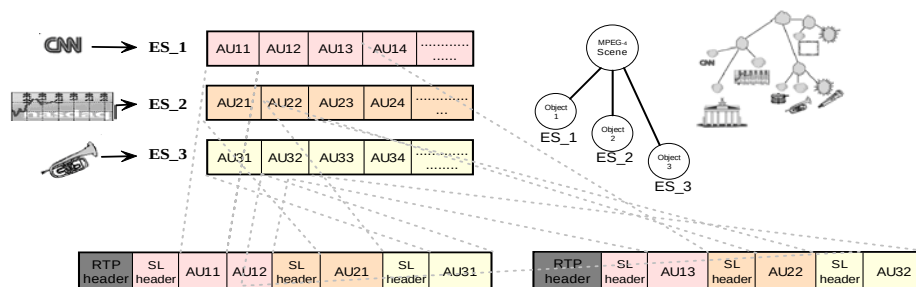


Figure 3.10 : Illustration du processus d'encapsulation dans RTP4mux

L'objectif de concaténer plusieurs AU dans le payload RTP étant dans le but de réduire l'overhead, d'où une meilleure exploitation de la bande passante. Tandis que la motivation d'encapsuler des AU, issues de différents flux élémentaires (ES), est la réduction des délais de mise en paquets. De plus, ce double multiplexage permet une meilleure robustesse aux pertes en réduisant les dépendances entre les paquets RTP adjacents.

Dans le protocole RTP4mux [50], le multiplexage (regroupement) d'un certain ensemble de flux élémentaires dans un même payload RTP s'effectue sur la base de :

- la nature des flux élémentaires (flux étroitement synchronisés ou ayant un même besoin en terme de champs d'en-tête du payload RTP). Par exemple : flux d'animation d'une bouche et le flux CELP de

parole associée, etc. Cette contrainte de synchronisation est aussi gérée par l'utilisation d'un en-tête SL réduit.

- les exigences en QoS des flux, c'est-à-dire réaliser une sorte d'agrégation des ES dans un même paquet RTP (ES ayant les mêmes besoins en QoS) pour faciliter le mapping de la QoS sur le réseau de transport sous-jacent. Ceci est particulièrement intéressant pour un réseau DiffServ.

Pour remédier aux problèmes d'overhead dûs à la redondance ou l'inutilité de certains champs de l'en-tête SL, RTP4mux adopte une encapsulation des flux SL avec une mise en correspondance « mapping » des champs redondants, mais surtout avec la possibilité de concaténer plusieurs AU dans le payload (dans la limite fixée par le path-MTU).

Ci-dessous les principales fonctions du protocole RTP4mux :

- **Fragmentation et encapsulation des flux élémentaires** : cette fonction prend en considération la nature du flux à transporter, en fragmentant ce dernier en entités indépendamment décodables. De plus, la partie encapsulation améliore la tolérance aux pertes au niveau application à travers le déploiement d'un mécanisme d'entrelacement des flux élémentaires.
- **Multiplexage et démultiplexage des flux élémentaires** : cette fonction de multiplexage est réalisée à 2 niveaux, et est destinée aux réseaux IP bas débit. D'une part, nous multiplexons plusieurs flux élémentaires sur une même session RTP. D'autre part, nous multiplexons des AU, appartenant à des ES différents, dans un même payload RTP.
- **Synchronisation des flux élémentaires** : cette fonction, également à la charge de RTP4mux, permet la synchronisation inter et intra flux élémentaires agrégés. Cette synchronisation repose sur le système MPEG-4 et évite, par conséquent, une synchronisation au niveau transport, beaucoup plus complexe et contraignante pour faire des agrégats.
- **Protection inégale aux erreurs** : cette protection inégale aux erreurs s'applique parfaitement sur les sources vidéo codées selon un modèle objet. Le taux de protection assigné à un objet dépend de son niveau relatif de priorité. Le rapport entre la bande passante utilisée et la distorsion induite par les pertes est le critère à optimiser.

3.3.3 Un Algorithme de contrôle du débit « TCP-friendly » pour les sources vidéo MPEG-4

La majorité des applications multimédia utilise la pile protocolaire RTP au dessus du protocole UDP/IP. Cependant, le protocole UDP n'offre aucun mécanisme de contrôle de congestion et il est inéquitable envers d'autres trafics. Le trafic Internet est dominé par des flux TCP, ce dernier utilise plusieurs mécanismes de

contrôle de congestion tels que : *AIMD* (*Additive Increase and Multiplicative Decrease*), *slow start*, *congestion avoidance*, etc. De ce fait, il est nécessaire que UDP réalise aussi un contrôle de congestion ou tout simplement soit compatible TCP (TCP-Friendly) [MAH 97].

L'idée d'un transport TCP-Friendly est d'émuler le fonctionnement du protocole TCP sans toutefois le dupliquer. Plusieurs algorithmes TCP-Friendly ont été développés dans le cadre du transport de la vidéo sur IP (voir section 3.2.1). L'IETF a récemment proposé l'algorithme TFRC [HAN 03] qui s'est imposé comme le meilleur protocole pour réaliser l'équité du trafic entre plusieurs flux. **Notre système de streaming vidéo utilise une extension TFRC pour supporter une adaptation des objets AVO aux variations de la bande passante du réseau.** Notre approche est basée sur l'algorithme suivant :

- Le client mesure les pertes rencontrées et retourne le taux d'erreur au serveur de streaming. Ce rapport d'erreurs est rendu possible grâce aux messages de contrôle RTCP (RR : Receiver Report) envoyés périodiquement par le client. En effet, chaque paquet RTP de données contient dans son en-tête des informations relatives aux timestamp et au numéro de séquence, il est donc possible pour le client de calculer les taux de pertes et les délais inter-arrivées des paquets.
- Les valeurs du taux de perte et du délai inter-arrivées des paquets sont utilisées par le module TFRC pour calculer le débit autorisé selon le modèle TFRC.
- Le serveur ajuste son débit en fonction du débit calculé en ajoutant/supprimant des objets AVO dans la scène en fonction de leur importance.

Le débit calculé par TFRC est obtenu par la formule suivante :

$$R_{TCP} \cong \frac{s}{RTT \sqrt{\frac{2bp}{3}} + t_{RTO} (3\sqrt{\frac{3bp}{8}}) p(1 + 32p^2)}$$

La valeur R_{TCP} représente le débit autorisé pour la transmission, s est la taille d'un paquet, RTT est la valeur du temps aller-retour, p est le taux de perte, t_{RTO} est le temps de retransmission et b est le nombre de paquets acquittés par un seul paquet d'acquiescement.

Afin de permettre au serveur d'ajuster dynamiquement son débit, nous avons modélisé la scène MPEG-4 par un ensemble d'objets AVO. Soit E , l'ensemble des objets MPEG-4 triés par rapport à leur importance dans la scène suivant le modèle de classification présenté précédemment :

$$E = \begin{cases} L_{1,1}, L_{1,2}, \dots, L_{1,m1} \\ L_{2,1}, L_{2,2}, \dots, L_{2,m2} \\ \dots \\ L_{n,1}, L_{n,2}, \dots, L_{n,mn} \end{cases}.$$

Avec $L_{x,y}$ correspondant à une couche MPEG-4 y d'un AVO x , ou une entité décodable par exemple une image

I, P, ou B. L'ensemble E est aussi noté $E = \{e_1, e_2, \dots, e_w\}$ avec $w = |E| = \sum_{j=1}^n m_j$

Le serveur vidéo ajoute un nouvel objet audio-visuel lorsque le débit estimé R_{TCP} excède le débit actuel de la session à la sortie du serveur. Supposons que le serveur est en train d'envoyer k entités (objet ou couche audiovisuel) à l'instant t_i . Le comportement du serveur à l'instant t_{i+1} doit obéir à la règle suivante :

$$\sum_{j=1}^{k+1} R_{i+1}(e_j) \leq R_{TCP}$$

Au niveau du client, les nouveaux objets reçus seront bufférisés et synchronisés pour être joués au même titre que l'ensemble des objets.

De la même manière, si le débit estimé R_{TCP} indique que le serveur consomme plus qu'il ne faut pour sa session, alors il doit réduire son débit en arrêtant le streaming des objets les moins importants, et ceci tant que la condition suivante est réalisée :

$$\sum_{j=1}^k R_{i+1}(e_j) > R_{TCP}$$

L'algorithme TFRC calcule le débit autorisé du serveur à chaque RTT. De ce fait, ajouter et supprimer des objets audiovisuels à chaque RTT peut produire des oscillations indésirables et une qualité faible au niveau du lecteur vidéo. Pour remédier à ce comportement oscillatoire, deux mécanismes ont été proposés.

Premièrement, une implémentation optimale de la moyenne exponentielle EWMA (Exponentially Weighted Moving Average) permet de détecter les situations hors contrôle de manière rapide. En effet, EWMA est utilisé pour répondre dynamiquement au changement de valeurs du RTT et des taux de perte et régularise ces valeurs pour traduire la réalité. Cela permet de régulariser le débit afin qu'il ne change pas de manière agressive.

Deuxièmement, un autre mécanisme fonctionne en collaboration avec le mécanisme précédent est implémenté. Il permet d'assurer la synchronisation des images vidéo envoyées par un lissage au niveau du **GOV** (Group of video object plane). Le serveur vidéo est autorisé à changer son débit seulement au début de chaque GOV. La moyenne des débits de toutes les images appartenant au même GOV est considérée dans le processus d'ajustement du débit.

Nous avons effectué des expérimentations sur une plateforme de test afin de valider l'efficacité du système proposé pour le streaming des services VoD. Les résultats obtenus en termes de critères objectifs (garanti de délai de bout-en-bout, réduction des taux pertes, amélioration du PSNR) et subjectifs (meilleure qualité visuelle, stabilité de visualisation) démontrent l'efficacité de nos propositions.

3.4 Conclusion

La demande de services vidéo sur les réseaux IP se confirme plus que jamais. Cependant, plusieurs caractéristiques de ces réseaux font que le déploiement à grande échelle de tels services présente un défi réel. Les services audiovisuels doivent faire face aux différentes variations de la bande passante et du délai de transfert, tout en restant robustes face aux pertes de paquets et aux erreurs de transmission.

Dans notre activité de recherche, nous avons examiné ces différents problèmes, et nous avons conçu et validé un système de diffusion vidéo adaptatif et intégré (« cross-layer »). Ce système résout certains problèmes induits par le transport de la vidéo en mode paquet sur les réseaux IP avec garantie statistique de la qualité de service (i.e. IP DiffServ).

L'originalité de notre contribution est de proposer un système adaptatif de diffusion vidéo respectant l'approche intégrée également appelée « cross-layer ». Les mécanismes proposés permettent de mieux contrôler la QoS des flux vidéo transportés sur les réseaux IP.

Conformément aux modèles architecturaux MPEG-4 et IP DiffServ, nos contributions s'articulent autour (1) d'un modèle de classification automatique d'objets audiovisuels intégré à la couche système MPEG-4, (2) d'un protocole robuste de fragmentation et d'encapsulation de flux vidéo sur RTP avec contrôle inégal d'erreurs appelé RTP4mux, et (3) un algorithme de contrôle de débits (TCP-Friendly) pour les flux MPEG-4 asservi aux variations de bande passante et aux caractéristiques des flux multimédia.

Finalement, pour favoriser une plus grande interopérabilité des services multimédia MPEG-4 avec les terminaux IP fixes et mobiles, nous avons également proposé, conçu et implémenté une architecture de passerelle de signalisation supportant les protocoles de contrôle de sessions IETF SIP (Session Initiation Protocol) et OSI MPEG-4 DMIF (Delivery Multimedia Integrated Framework). Cette passerelle multimédia de signalisation décrite dans [AHM 02] permet d'assurer l'interopérabilité des services MPEG-4 entre les réseaux DVB et IP, la résolution d'adresse et la négociation des capacités des terminaux IP fixes ou mobiles.

3.5 Mes Publications Associées

- [1] A. Mehaoua, Toufik Ahmed, H. Asgari, M. Sidibe, A. Nafaa, G. Kormentzas, T. Kourtis, "Service-driven inter-domain QoS Monitoring System for Large-scale IP and DVB Networks", in Computer

Communication Journal, SI on Monitoring and Measurements of IP Networks, Vol. 29, No. 10, pp. 1687-1695, June 2006.

- [2] Toufik Ahmed, Raouf Boutaba and Ahmed Mehaoua, "A Measurement-based Approach for Dynamic QoS Adaptation in DiffServ Networks", in Computer Communications Journal, Special issue on End-to-End Quality of Service Differentiation, Elsevier Editor, VOL. 28, Issue 18, pp. 2020-2033, November 2005.
- [3] Toufik Ahmed, Ahmed Mehaoua, Raouf Boutaba, and Youssef Iraqi "Adaptive Video Streaming Over Next Generation Networks: A Cross-layer Approach" in IEEE Journal on Selected Areas in Communications (J-SAC), Special issue on Intelligent Services and Applications in Next Generation Networks, Vol. 23, N°. 2, pp. 385 - 401 February. 2005.
- [4] Vincent Lecuire, Toufik Ahmed, et Ahmed Mehaoua, "Protection des Flux Audiovisuels MPEG-4 Transportés sur Internet", dans la revue Technique et Science Informatique: Réseaux et Protocoles, TSI vol. 23, n° 5-6/2004, pp. 623-650, Ed. Hermes, ISSN 0752-4072, Mai 2004.
- [5] Toufik Ahmed, Ahmed Mehaoua, Raouf Boutaba, Youssef Iraqi, "IP Video Streaming With Fine-Grained TCP-Friendly Rate Adaptation", dans Lecture Notes in Computer Science (LNCS), "Management of Multimedia Networks and Services" (MMNS 2003), Springer-Verlag Editor, vol. 2839, pp. 18-31, 2003.
- [6] Toufik Ahmed, Ahmed Mehaoua et Vincent Lecuire "Streaming MPEG-4 Audiovisual Objects Using TCP-Friendly Rate Control And Unequal Error Protection", dans IEEE International Conference On Multimedia & Expo (IEEE Multimedia'03), Ed. IEEE Press, Volume 2, pp. 317-320, Baltimore, July 2003.
- [7] A. Nafaa, Toufik Ahmed, Y. Hadjadj Aoul et A. Mehaoua "RTP4Mux: A Novel MPEG-4 RTP Payload for Multicast Video Communications Over Wireless IP" dans IEEE Packet Video Workshop (Packet Video'03), April 28-29, Nantes, France, April 2003.
- [8] Toufik Ahmed, A. Mehaoua, R. Boutaba, "Interworking Between SIP and MPEG-4 DMIF for Heterogeneous IP Videoconferencing", dans IEEE International Conference on Communications (ICC'02), Vol. 4, pp. 2469-2473, IEEE Press, NewYork, Avril 2002.
- [9] Toufik Ahmed, A. Mehaoua et G. Buridant "Implementing MPEG-4 Video On Demand Over IP Differentiated Services" dans IEEE International Conference on Global Communications (Globecom'01), Ed. IEEE Press, Volume 4, pp. 2489-2493, San Antonio, Texas, USA, Novembre 2001.
- [10] Toufik Ahmed, Ahmed Mehaoua et R. Boutaba "Un modèle de contrôle de service de vidéoconférence IP impliquant des terminaux multimédia SIP et MPEG-4" dans CFIP'02, Montréal Canada, Mai 2002.

3.6 Bibliographie

- [AHM 02] T. Ahmed A. Mehaoua, R. Boutaba, "Interworking Between SIP and MPEG-4 DMIF for Heterogeneous IP Videoconferencing", dans IEEE International Conference on Communications (**ICC'02**), Vol. 4, pp. 2469-2473, IEEE Press, New York, Avril 2002.
- [BAS 02] Basso, Civanlar, P. Gentric, Herpel, Lifshitz, Lim, Perkins, Van Der Meer, "RTP Payload Format for MPEG-4 Streams", the Internet Engineering Task Force (IETF), July 20, 2001, work in progress, expires January 2002.

- [CUR 01] D.Curet, E.Gouleau, S.Relier, C.Roux, P.Clement, G.Cherry, "RTP Payload Format for MPEG-4 FlexMultiplexed Streams", Internet Draft, IETF, Work in progress, November 2001.
- [HAN 98] M. Handley, V. Jacobson, "RFC: 2327 SDP Session Description Protocol", Request for Comments, IETF, Apr. 1998.
- [HAN 03] M.Handley, S.Floyd, J.Padhye, and J.Widmer "RFC 3448, TCP Friendly Rate Control (TFRC): Protocol Specification" Request for Comments, IETF, Jan. 2003.
- [ISO-IEC 14496.1 98] ISO/IEC 14496-1 "Coding of audio-visual objects, Part 1: Systems", final committee draft, May 1998.
- [ISO-IEC 14496.2 98] ISO/IEC 14496-2 "Coding of audio-visual objects, Part 2: Visual", final committee draft, May 1998.
- [ISO-IEC 14496.3 98] ISO/IEC 14496-3 "Coding of audio-visual objects, Part 3: Audio", final committee draft, May 1998.
- [ADA 03] Adam H. Li, et al., "An RTP Payload Format for Generic FEC with Uneven Level Protection", *Internet draft*, draft-ietf-avt-ulp-07.txt, work in progress, expires: May 4, 2003.
- [AMI 95] Amir E., McCanne S., and Zhang H., "An Application Level Video Gateway", *In Proc. ACM Multimedia*, 1995.
- [AMI 98] Amir E., McCanne S., and Katz R., "An Active Service Framework and its application to real-time multimedia transcoding", in SIGCOMM, symposium on communications architectures and protocols, Septembre 1998.
- [FLO 00] Floyd S., Handley M., Padhye J., and Widmer J., "Equation-based congestion control for unicast applications", In Proc. of ACM SIGCOMM, pp. 43–56, 2000.
- [FUJ 01A] Fujimoto K., Ata S., and Murata M., "Statistical analysis of packet delays in the Internet and its application to playout control for streaming applications", IEICE Transactions on Communications, vol. E84-B, pp. 1504–1512, June 2001.
- [FUJ 01B] Fujimoto K., Ata S., and Murata M., "Playout control for streaming applications by statistical delay analysis," in Proceedings of IEEE International Conference on Communications (ICC 2001), vol. 8, (Helsinki), pp. 2337–2342, June 2001.
- [HAN 03] Handley M., Floyd S., Padhye J., and Widmer J., "RFC 3448, TCP Friendly Rate Control (TFRC): Protocol Specification", Request for Comments, IETF, Jan. 2003.
- [JAC 88] Jacobson V., "Congestion Avoidance and Control", ACM SIGCOMM, August 1988.
- [MAH 97] Mahdavi M., and Floyd S., "TCP-Friendly Unicast Rate-Based Flow Control", Technical note sent to the end2end-interest mailing list, January 8, 1997.
- [KAL 02] Kalman M., Steinbach E., and Girod B., "Adaptive playout for real-time media streaming, in Proc. IEEE Int. Symp. on Circ. and Syst. May 2002.
- [LI 98] Li X., Paul S., and Ammar M., "Layered Video Multicast with Retransmission (LVMR): Hierarchical Rate Control Schemes", In Proc. of Infocom 98, San Francisco, CA, March 1998.

- [LIE 03] Liebl G., Wagner M., Pandel J. and Weng W., "An RTP Payload Format for Erasure-Resilient Transmission of Progressive Multimedia Streams", Internet draft, draft-ietf-avt-uxp-04.txt, work in progress, expires: May 2003.
- [LIN 83] Lin S., Costello D.J., "Error Control Coding: Fundamentals and Applications", Prentice Hall, 1983.
- [MAR 96] Marasli R., Amer P. D., and Conrad P. T., "Retransmission-Based Partially Reliable Transport Service: An Analytic Model", INFOCOM, pp. 621-629, 1996.
- [REA 99] Real Network, "RealSystem G2: Management and Control of Streaming Media Over Corporate Networks RealVideo", available at: <http://docs.real.com/docs/devzone/g2ctrlandmgmt.pdf>, Mar. 1999.
- [REJ 99] Rejaie R., Handley M., and Estrin D., "RAP An End-to-End Congestion Control Mechanism for Realtime Streams in the Internet". In Proc. IEEE Infocom '99, New York, NY, March 1999.
- [RHE 00] Rhee I. and Joshi S., "Error recovery for interactive video transmission over the Internet", IEEE J. Selected Area Comm., vol. 18, no. 6, pp. 1033-49, June 2000.
- [SIS 98] Sisalem D., and Schulzrinne H., "The Loss-Delay Adjustment Algorithm: A TCP-friendly Adaptation Scheme, Network and Operating System Support for Digital Audio and Video", in NOSSDAV, Cambridge, UK, 8-10, July 1998.
- [STE 01] Steinbach E. G., Farber N., and Girod B., "Adaptive Playout for Low Latency Video Streaming," Proc. International Conference on Image Processing (ICIP-01), Oct. 2001.
- [SUN 02a] Sun X., Li S., Wu F., Shen G., and Gao W., "Efficient and flexible drift-free video bitstream switching at predictive frames", in ICME'02, 2002.
- [SUN 02B] Sun X., Wu F., Li S., Gao W., and Zhang Y. Q., "Seamless Switching of Scalable Video Bitstreams for Efficient Streaming", IEEE International Symposium on Circuits and Systems, ISCAS 2002, Scottsdale, Arizona, USA, 26-29 May, 2002.
- [VAN 04] J. Van Der Meer, D. Mackie, V. Swaminathan, D. Singer, P. Gentric, "RTP Payload Format for Transport of MPEG-4 Elementary Streams", Internet Draft, IETF, draft-ietf-avt-mpeg4-simple-08.txt, Work in progress, August 2003, Expires February 2004.
- [YOS 00] Yoshihiro K. al., "RFC: 3016 RTP Payload Format for MPEG-4 Audio/Visual Streams", the Internet Engineering Task Force (IETF), November 2000.

4 Transport et contrôle de la qualité des services vidéo sur réseaux sans-fil

4.1 Introduction

Les réseaux de communication sans fil et mobiles de type 802.11 connaissent actuellement un succès incontestable. Ceci s'explique principalement par la facilité avec laquelle on peut déployer un réseau sans fil 802.11 et par la multiplication des terminaux mobiles supportant cette technologie (PC portables, téléphones mobiles, PDA communicants, Smart Phones, Tablet PC, etc.). Les avancées dans les techniques de compression numérique des séquences vidéo (H.264 AVC/SVC) et l'incessante croissance de la capacité des réseaux 802.11 ont favorisé l'émergence de services multimédia sur les réseaux sans fil. Ainsi, de plus en plus d'opérateurs réseaux proposent des « hotspot » d'accès Internet couplés à des expérimentations de diffusion des services multimédia, tel que la diffusion de la télévision et des séquences vidéo. Cependant, la QoS reste le premier verrou technologique à lever avant la généralisation de ce genre de service. La fiabilité des communications représente un des principaux obstacles à la distribution des flux vidéo dans un réseau 802.11. Les expérimentations que nous avons menées ont démontré que la distribution des erreurs et des pertes dans un environnement 802.11 suit un processus aléatoire dépendant de plusieurs facteurs difficiles à identifier et à quantifier. Il est pratiquement impossible de reproduire le même comportement de deux expérimentations s'exécutant sous les mêmes conditions. D'autant plus que la mobilité des utilisateurs agit sur la capacité de réception et influence directement le comportement d'autres utilisateurs. Les spécificités du lien 802.11 en termes de partage de bande passante et de distribution de pertes de paquets sont deux phénomènes qui doivent être analysés afin de mieux contrôler la QoS pour les services vidéo sans fil.

C'est l'objet de ma deuxième contribution dans laquelle je me suis, en conséquence, intéressé au transport et contrôle de la qualité de service vidéo sur les réseaux sans fil de type 802.11. Ce travail a été mené dans le cadre de la thèse de doctorat de M. Ismail Djama. Ainsi, nous avons exploité des techniques Cross-Layer (XL) pour l'optimisation des ressources et le contrôle de QoS pour les services vidéo sans fil.

4.2 État de l'art

Actuellement, l'architecture d'Internet est basée sur le modèle protocolaire TCP/IP. Cette architecture est caractérisée par sa structure modulaire mettant en œuvre plusieurs couches distinctes totalement isolées qui communiquent à travers des interfaces bien définies. Ainsi, chaque couche dépend de la fonction de service de la

couche sous-jacente. Pour fournir un service donné, la couche inférieure utilise l'encapsulation pour placer les données de la couche supérieure dans son champ de données. Ainsi, sur une vue verticale, chaque couche augmente les données utiles par des en-têtes contenant des informations de service spécifiques à chaque couche dans la pile TCP/IP. D'autre part, sur une vue horizontale et au niveau des systèmes distants/communicants, les couches du même niveau se coordonnent et communiquent à l'aide des champs d'en-tête pour implémenter et mettre en œuvre des services avancés de contrôle de congestion (acquittement et rapports de pertes), contrôle d'erreurs (Checksum et CRC), etc.

Au fil de l'évolution de l'Internet, plusieurs mécanismes et protocoles ont été ajoutés aux différentes couches pour implémenter des besoins majeurs en termes de qualité de service (QoS), de sécurité, de mobilité, etc. Cependant, plusieurs mécanismes ont été dupliqués inutilement sur plusieurs couches ou inversement ont été localisés sur quelques couches protocolaires. Un exemple typique est le mécanisme de contrôle d'erreur qui se fait à la fois au niveau de la couche « liaison de données », mais aussi au niveau de la couche « transport » pour palier aux mêmes problèmes de fiabilité mais à des échelles temporelles et spatiales différentes : le CRC et la retransmission sont deux mécanismes utilisés pour détecter et retransmettre les paquets perdus. Ces deux mécanismes, comme celui de l'adaptation du codage et modulation, se font d'une manière complètement indépendante sur la base de contexte disponible au niveau de leur couche (en-têtes).

Ceci rend la continuité d'un mécanisme à travers les couches réseaux (exemple : continuité de QoS) un problème important pour l'efficacité des communications et l'optimisation des ressources. Une optimisation globale des différentes couches permet de résoudre ce problème. Récemment, une nouvelle approche d'optimisation des protocoles de communication réseaux a vu le jour avec l'introduction du modèle reposant sur une interaction étroite entre couches appelé **Cross-Layer (XL)**. Ce modèle basé sur le concept « **Think globally, act locally** » permet de déterminer la meilleure configuration du système pour la réalisation d'un objectif. L'approche Cross-Layer (XL) augmente l'interaction entre les différentes couches pour exploiter les caractéristiques inhérentes du réseau afin de maximiser l'utilité du système (par exemple maximiser la QoS) et pour réduire le coût (par exemple, optimiser la durée de vie de la batterie). Plus concrètement, les mécanismes de contrôle d'erreurs peuvent être coordonnés tout au long de la pile TCP/IP pour produire une meilleure efficacité en termes de précision de l'adaptation et améliorer, par la même, l'utilisation de la bande passante.

Le concept Cross-Layer (XL) a récemment émergé avec la prolifération des réseaux sans fil, mais aussi, avec l'augmentation du nombre de clients hétérogènes connectés à travers différents types de réseaux d'accès. Dans le modèle XL, il est souvent plus judicieux de s'affranchir de certains mécanismes TCP/IP futiles pour réduire le volume du trafic de contrôle (overhead) et maximiser ainsi l'efficacité globale du système communicant. Ceci peut s'avérer particulièrement indispensable dans les réseaux sans-fil où la taille moyenne des paquets au niveau 2 (Lien) peut être contrôlée par des mécanismes de fragmentation pour faire face aux contraintes de l'interface air (i.e., interférence co-canal, fading, shadowing, bruit, etc.) et maximiser l'efficacité des différentes couches

physique (codages et modulation). Il est communément accepté que les paquets de petites tailles soient plus robustes dans un environnement sans fil qui présente des taux d'erreurs élevés, bien que les paquets de plus grande taille favorisent l'utilisation de la bande passante en réduisant l'overhead impliqué par les en-têtes et le délai de bout-en-bout.

La raison initiale qui a motivé l'émergence de l'approche cross-layer (XL) est principalement dérivée des raisons suivantes :

- Les réseaux sans fil sont caractérisés par des erreurs de transmission, des schémas de modulation et de codage variés et dynamiques, des interférences co-canal, etc. Pour remédier à ces problèmes, les couches protocolaires doivent coopérer et collaborer pour une meilleure efficacité de la transmission. Le service vidéo, dans ce cas, doit s'adapter au changement effectué au niveau des couches inférieures. Cette adaptation doit être optimale, efficace, et doit répondre aux besoins des utilisateurs pour l'optimisation globale du service fourni.
- L'efficacité globale d'une transmission sur un réseau sans fil a été et continu d'être l'objet de plusieurs travaux de recherche. L'optimisation des différentes couches de transmission est le premier objectif pour considérer le modèle cross-layer (XL). Il existe plusieurs champs optionnels le long de toutes les couches TCP/IP qui peuvent être exploités pour rendre possible une communication/coordination entre des mécanismes situés dans différentes couches afin d'accomplir un objectif commun fixé par l'application.
- L'hétérogénéité des applications et des réseaux exigent des mécanismes plus rigoureux d'adaptation. En particulier, dans le contexte des services multimédia, l'adaptation du contenu est nécessaire pour remédier au problème de l'hétérogénéité. En effet, les réseaux 802.11 ont introduit d'incontestables avancées dans l'adaptation au niveau 2 pour servir de manière différente et plus efficace divers récepteurs (selon leurs conditions de réception respectives) sans pour autant compromettre l'efficacité totale du système (exemple : adaptation par le récepteur du débit physique du lien 802.11). Cette adaptation exige une plus grande « *cognitivité* » de la part des couches supérieures en adaptant leur débit, leur format de codage, leur contrôle d'erreur applicatif (FEC), etc. Ainsi, si la couche lien s'adapte et ajuste le débit de transmission, les applications doivent aussi s'auto-discipliner pour éviter toute situation contre-productive.
- Les mécanismes cross-layer (XL) permettent d'offrir une continuité de QoS à travers les couches protocolaires. Ceci passe par une plus grande coopération entre les couches adjacentes afin de traduire les exigences de l'application multimédia jusqu'à la couche lien. À titre d'exemple, il serait possible d'envisager un mécanisme DiffServ au niveau IP qui pourrait consulter des informations insérées par des couches supérieurs (par exemple RTP) pour déterminer le type de trafic (voix ou vidéo par exemple) et

le privilégier dans sa politique de scheduling (ordonnancement). En effet, rien n'empêche les architectes de solutions réseaux de casser le paradigme de la séparation et d'indépendance des couches TCP/IP ; l'efficacité étant le seul objectif qui prévaut.

4.2.1 Introduction au concept Cross-Layer (XL)

La multiplication des mécanismes de QoS et des techniques d'adaptation sur les différentes couches du modèle TCP/IP a engendrée plusieurs problématiques causées principalement par le principe d'isolation des couches. Ces problématiques peuvent être regroupées en trois grandes classes :

- La redondance : elle est causée par la duplication d'un mécanisme sur plusieurs couches. Nous pouvons prendre l'exemple de la retransmission dans TCP et dans la couche MAC 802.11. Même si les deux mécanismes opèrent sur des échelles temporelles et spatiales différentes, il faut une optimisation globale du système pour maximiser la QoS.
- L'annulation : certains mécanismes introduits sur des couches supérieures ne sont pas respectés par des couches inférieures ce qui provoque leurs annulations. Par exemple, la priorité introduite au niveau IP grâce aux classes de services DiffServ n'est pas forcément assurée au niveau de la couche MAC 802.11 si celle-ci n'implémente pas le protocole 802.11e ou équivalent. Dans ce cas de figure, la continuité de QoS ne peut se faire si un maillon de la chaîne brise ce principe.
- La contradiction : dans certain cas extrêmes, l'effet de deux mécanismes présents sur deux couches distinctes s'avère opposé. Par exemple, l'utilisation d'UDP au niveau transport pour éviter les retransmissions et l'utilisation de la retransmission au niveau de la couche MAC 802.11 peut être vu comme contradictoire. La contradiction signifie aussi l'annulation du mécanisme. Par exemple, dans le cadre des réseaux IP DiffServ, la contradiction entre domaines adjacents dans le marquage des flux peut dégrader sensiblement la QoS. Un flux marqué initialement avec une priorité EF dans un domaine, ne doit pas se trouver marqué avec une priorité BE dans un autre domaine.

Afin d'apporter une solution à cette problématique et optimiser les performances des systèmes communicants, le concept XL est utilisé dans notre démarche de conception de nouveaux mécanismes de contrôle de QoS pour les services vidéo transportés sur les réseaux sans fil. En effet, le concept XL autorise la violation de la structure protocolaire pour contrer le problème d'isolation des couches. Il nous permet une meilleure coordination des mécanismes existants ou de créer de nouveaux mécanismes non possibles avec le modèle classique.

Comme tout nouveau concept, il est très difficile de trouver ou de proposer une définition exacte du « Cross-layer ». Même au niveau de la terminologie, nous trouvons dans la littérature plusieurs variantes : la conception cross-layer, l'adaptation cross-layer, l'optimisation cross-layer, le retour d'information cross-layer. Dans [SRI 05] [VAN 05], les auteurs ont proposé une définition générique qui peut englober toutes les

techniques et tous les mécanismes cross layer qui existent actuellement. Ainsi, le design cross-layer est défini comme étant : « la conception d'un protocole en violation avec l'architecture en couche de référence est une conception cross-layer à l'égard de cette architecture ». Le terme violation englobe :

- La définition de nouvelles interfaces entre les couches, pour passer / recevoir des paramètres de services.
- La redéfinition des limites des couches.
- La conception d'un protocole sur une couche en se basant sur la conception d'un autre protocole sur une autre couche.
- La configuration commune des paramètres à travers les couches.

C'est ainsi que le concept Cross-Layer (XL) permet la définition de protocoles ou de mécanismes qui ne respectent pas l'isolation des couches du modèle OSI (respectivement TCP/IP). Il autorise donc la communication entre deux ou plusieurs couches, adjacentes ou non, dans le but d'améliorer les performances globales du système.

4.2.2 Signalisation dans les architectures Cross-Layer

Dans le modèle TCP/IP, la signalisation des données est assurée par les champs d'en-tête des différentes couches. Cependant, il n'est pas permis à une couche inférieure d'aller chercher les valeurs des champs se trouvant à une distance de plus de deux couches. C'est le principe même d'isolation des couches qui impose cette restriction. L'unité de donnée envoyée par la couche supérieure (en-tête et payload) est reçue comme élément obscur qui doit être encapsulé et renvoyé à la couche inférieure. Dans le modèle cross-layer la signalisation des données peut se faire suivant différents schémas. Parmi toutes les architectures « cross-layer » proposées dans la littérature, deux modèles de communications peuvent être distingués [SRI 05] [SED 06]. Le premier modèle consiste à faire une communication directe entre les couches. C'est un modèle dans lequel la couche qui a besoin d'une information particulière, accède et récupère cette information grâce à la nouvelle interface définie et prévue à cet effet. Le deuxième modèle repose sur l'utilisation d'une base de données partagée entre les couches (shared database) qui joue le rôle d'un hub de données. Les informations relatives à chaque couche peuvent être concentrées au niveau du hub pour les mettre à disposition d'autres couches. Ce modèle est plus simple à mettre en œuvre mais son efficacité est réduite par rapport au premier modèle.

Les sections suivantes détaillent d'avantage ces deux modèles de signalisation dans les architectures cross-layer.

4.2.2.1 Signalisation directe entre les couches

La communication directe entre les couches est le modèle le plus utilisé par les architectures « cross-layer ». Il permet à une couche d'accéder directement aux paramètres et aux variables d'une autre couche sans passer par un intermédiaire. Cette communication peut être « in-band » en utilisant l'en-tête des protocoles qui sont

déployés actuellement, par exemple, la couche IP accède aux champs de l'en-tête TCP « ECN » pour s'informer d'une congestion dans le réseau. Dans certain cas, des extensions d'en-tête sont nécessaires pour faire passer des informations supplémentaires. La communication directe peut être « out-of-band » en utilisant un nouveau protocole dédié. Ainsi, le protocole de signalisation CLASS (Cross-Layer Signaling Shortcuts), présenté dans [WAN 03], est un parfait exemple de ce type de protocole. La communication « out-of-band » peut s'effectuer aussi en définissant de nouvelles interfaces (API) au niveau d'une couche qui seront utilisées directement par d'autres couches pour récupérer et configurer des paramètres de fonctionnement.

4.2.2.2 Signalisation via une base de données partagée

Plusieurs architectures « cross-layer » [WIN 06] [KHA 06] proposent l'utilisation d'une base de données partagée ou un hub afin de stocker et de récupérer des paramètres de configuration de certaines couches. Cette base est accessible par toutes les couches qui peuvent, ainsi, s'informer de l'état des autres couches ou récupérer des paramètres de configurations nécessaires à leur fonctionnement interne. La base de données est aussi accessible par un système d'optimisation responsable d'initialiser les paramètres avec les valeurs adéquates. Cependant, pour mettre en œuvre cette base de données, il faut répondre à plusieurs contraintes conceptuelles comme la localisation de la base de données (sur les nœuds communicants ou sur un nœud indépendant), le type de communication entre les couches et la base de données partagée, ainsi que le protocole de communication utilisé.

4.2.3 Les interactions Cross-Layer

Actuellement, plusieurs techniques « cross-layer » sont proposées pour améliorer les performances des applications utilisant principalement les réseaux sans fil. Historiquement, ces mécanismes étaient limités à des interactions simples entre la couche physique et la couche liaison de données. De plus, les mécanismes proposés étaient indépendants et visaient essentiellement à l'amélioration d'une imperfection précise. Ce concept est en phase de maturation grâce notamment à la modélisation des interactions pouvant exister entre couches et à l'apparition d'approches qui tentent de faire collaborer plusieurs couches, prenant en charge plusieurs paramètres, pour une optimisation globale. Trois grandes familles d'approches existent :

- L'approche ascendante (Bottom-up) : dans cette approche, les couches supérieures optimisent leurs mécanismes en fonctions des paramètres (conditions) des couches inférieures. Par exemple, le débit vidéo à la sortie d'un serveur peut varier en fonction des taux de pertes réseaux.
- L'approche descendante (Top-down): dans cette approche, les couches supérieures décident des paramètres de configuration des couches inférieures. De la même manière, les couches inférieures peuvent considérer certaines spécificités de niveau applicatif pour exécuter leurs traitements. Par

exemple, un serveur vidéo peut demander à la couche IP de faire un marquage DiffServ de son trafic pour refléter la priorité des images I, P et B.

- L'approche mixte (Integrated) : Cette approche exploite les deux approches précédentes dans une même architecture afin de trouver la meilleure configuration et optimisation possible pour un fonctionnement optimal du système.

4.2.4 Les projets européens traitant la problématique Cross-Layer

Plusieurs projets financés par la commission européenne ont traité, ou traitent, la problématique du « cross-layer » en étudiant ce nouveau concept et en proposant de nouvelles interactions « cross-layer » afin d'améliorer les performances de transmission. Nous présentons, ci-dessous, une brève description de quelques uns de ces projets.

Le projet 4MORE [4MORE] s'inscrit dans le développement de systèmes 4G, la nouvelle génération des communications mobiles. La vision européenne de la 4G est un nouveau système basé sur IP et offrant tous les services, à n'importe quel moment en utilisant n'importe quel terminal. Pour cela, la 4G doit offrir des débits variants entre 2 et 100 Mbps couvrant plusieurs environnements mobiles (véhicules, piétons) et fixes (intérieures et extérieures) dans une plage de fréquences de 50 – 100 MHz. La technologie de transmission utilisée pour atteindre cet objectif est le MC-CDMA qui est déjà adoptée par les Japonais. L'objectif du projet 4MORE est d'étudier, développer, intégrer et valider un système sur puce, rentable et de faible consommation, pour les terminaux mobiles utilisant un système de transmission multi-antennes basé sur MC-CDMA. Ce système doit intégrer de nouvelles interactions entre la couche physique (MC-CDMA) et les algorithmes de la couche MAC.

L'objectif du projet PHOENIX [PHO] est le développement d'un lien de communication de bout-en-bout basé sur un réseau sans fil optimisé. Pour cela, le projet propose une communication entre le monde applicatif (le codage source, le cryptage, etc.) et le monde des transmissions (le codage canal, la modulation) à travers le monde réseau, à savoir le protocole IPv6. Pour atteindre cet objectif, le projet se focalise sur trois axes principaux. Le premier axe se charge du développement d'un schéma innovant pour une optimisation conjointe du codage source et canal (JSCC : Joint Source Channel Coding). Cela inclut la définition de nouveaux codages source et canal et l'adaptation des codages qui existent actuellement. Le deuxième axe a pour objectif de définir de nouvelles stratégies d'adaptation efficaces qui prennent en considération des paramètres réels (présence de cryptage, le nombre de saut sans fil, etc.). Enfin, le troisième axe définit une nouvelle architecture réseau basée sur l'optimisation conjointe des couches protocolaires.

Le projet ENTHRONE II [ENT] propose une solution globale pour la gestion de la chaîne de distribution des services audiovisuels. Cette solution englobe la protection du contenu, la distribution à travers les NGNs et la

réception au niveau de l'utilisateur final. Le but de ce projet n'est pas d'unifier ou d'imposer une stratégie mais plutôt d'harmoniser les différentes fonctionnalités afin de supporter une qualité de service de bout-en-bout à travers des réseaux hétérogènes distribuant des services IP multimédia pour divers utilisateurs. Pour atteindre cet objectif, le projet se base sur une architecture de gestion ouverte et décentralisée de la chaîne de distribution de bout-en-bout. Le modèle MPEG-21 est utilisé comme un support commun pour l'implémentation et la gestion des fonctionnalités de l'architecture. L'adaptation MPEG-21 « cross-layer », qui a pour objectif d'adapter les contenus multimédias pour maintenir la QoS, est considérée comme un élément principal de l'architecture ENTHRON. Cette adaptation prend en considération plusieurs paramètres allant de la couche application (qualité objective du flux multimédia) jusqu'à la couche physique (débit physique dans les réseaux sans fil, niveau du signal) en passant par la couche réseau (latence, taux de pertes). **Une partie des travaux présentés dans ce chapitre s'inscrit dans le cadre de ce projet européen.**

4.2.5 Les réseaux WLAN 802.11

4.2.5.1 Niveau Physique 802.11

Le niveau physique 802.11 regroupe l'ensemble des règles et des procédures qui permettent d'acheminer des éléments binaires sur un support physique entre l'émetteur et le récepteur. Ce niveau prend en charge les aspects de codage et de modulation des signaux numériques électriques et optiques. La couche physique des réseaux sans fil est définie par la famille des normes IEEE 802.11, 802.11a/b/g et bientôt 802.11n. Ces différents standards décrivent des techniques d'occupation du spectre et des techniques de modulation du signal sur des bandes de fréquences bien déterminées. La technique de modulation définit le symbole qui sera utilisé au niveau du signal pour représenter les données, ainsi que le nombre de bits codés par symbole.

Les standards IEEE 802.11 introduisent plusieurs types de modulation et de taux de codage dont la combinaison fait varier le débit physique de 1 Mbits/s à 54 Mbits/s. Les débits élevés, comparés aux débits faibles, utilisent des modulations complexes qui placent plus de bits dans un intervalle de temps. Les modulations complexes sont plus sensibles aux bruits et nécessitent un certain niveau du signal pour leur décodage. Par conséquent, ceci limite leur portée. Inversement, les modulations simples qui fournissent de faibles débits donnent une plus grande portée puisqu'elles sont plus robustes et nécessitent un signal moins puissant pour leurs décodages.

Ainsi, la couverture d'un réseau sans fil 802.11 sera large mais fournira moins de débit si les modulations simples sont utilisées et cette couverture sera limitée mais offrira plus de débit si les modulations complexes sont utilisées. Les couches physiques de 802.11 se sont contentées de définir les modulations et les codages pouvant être utilisés sans toutefois définir les mécanismes et les algorithmes qui permettent de mieux contrôler ce débit et surtout de basculer d'un débit vers un autre.

4.2.5.2 Algorithme de contrôle de débit physique 802.11

Les algorithmes de contrôle du débit physique agissent sur la modulation et le codage utilisés par le niveau physique 802.11. Ces algorithmes appelés RCA (Rate Control Algorithm) permettent de déterminer le débit physique adéquat entre deux stations 802.11 pour minimiser les pertes et optimiser la communication. La différence entre plusieurs RCA réside dans la réactivité de l'algorithme, dans l'interprétation de l'état du canal et dans la politique utilisée pour le basculement du débit. Deux principaux modes sont utilisés dans l'interprétation de l'état du canal dans les algorithmes RCA. Le premier mode utilise des statistiques d'émission/réception des trames MAC, comme par exemple le taux d'erreur d'émission de trames, le nombre de retransmission, etc. Cela pour déterminer le débit physique entre deux stations. Des exemples d'implémentations sont donnés dans [LEC04] [KAM 97]. Ce mode permet une adaptation du débit à long terme mais sa réaction est limitée aux changements rapides de l'état du canal. Le deuxième mode utilise le niveau du signal SNR. Les RCA basés sur ce mode exploitent des techniques d'estimation de la puissance du signal perçue par le récepteur pour déterminer le meilleur débit physique. Des exemples d'implémentations sont donnés dans [PAV 03] [HOL 01]. Ce mode réagit rapidement aux changements de l'état du canal, mais il manque de fiabilité à cause de l'estimation approximative du SNR. Finalement, certains RCA, appelés RCA hybrides, se basent sur les deux modes à savoir statistique et SNR. Un exemple est donné dans [HAR].

Parce que le débit physique 802.11 est déterminé par des algorithmes au niveau de la couche liaison, il est d'une importance majeure pour les applications vidéo de s'adapter dynamiquement à ce changement de débit pour éviter toute situation contre productive dans laquelle les paquets vidéo sont supprimés au niveau de la file d'attente MAC avant leurs transmissions. Par conséquent, l'adaptation vidéo doit se faire au cours de la session et doit être dynamique pour refléter la quantité des ressources disponibles dans le canal physique en tenant compte du partage de ce canal par plusieurs utilisateurs. Ce contrôle de QoS rend l'application vidéo consciente (*aware*) de son environnement.

4.3 Contributions

Nous avons conçu et développé plusieurs mécanismes permettant le contrôle de la QoS pour le transport des flux vidéo temps réel sur les réseaux sans fil 802.11. Notre architecture appelée MIM (Meet In the Middle) permet de mieux contrôler les couches protocolaires par l'exécution des mécanismes d'adaptation cross-layer suivantes :

- Adaptation du débit de la transmission vidéo pour refléter le débit physique. En effet, permettre à l'application vidéo d'être consciente de son environnement de transport sous-jacent est essentiel pour la fourniture d'un service avec QoS. Ce mécanisme d'adaptation du débit répond au besoin de l'application de diffusion vidéo de refléter la quantité des ressources disponibles dans le canal de

transmission sans-fil afin de mieux servir les utilisateurs finaux. Pour ce faire, nous avons conçu un modèle mathématique permettant d'estimer le débit d'une station en fonction des débits des autres stations pour permettre un partage équitable de la bande passante et minimiser par la même les pertes de paquets durant la transmission.

- Adaptation conjointe FEC/débit pour contrôler la fluctuation permanente de la qualité du lien 802.11. Les erreurs et les pertes qui subsistent sur le lien 802.11 ne peuvent être contrôlées sans l'utilisation de codes correcteurs d'erreurs et de mécanismes de type FEC (Forwarding Error Correction). Ce mécanisme permet de mieux contrôler les erreurs de transmission sans fil sans toutefois provoquer plus d'erreurs dues au fait de l'augmentation du débit de la source.
- Fragmentation adaptative au niveau MAC 802.11. En effet, la probabilité de rejet d'une trame est proportionnelle à la taille de celle-ci. Plus la trame à transmettre sera grande, plus la probabilité, d'avoir une erreur binaire sur cette trame, sera grande aussi. Nous avons prouvé par des expérimentations que la fragmentation au niveau MAC est plus intéressante que la fragmentation au niveau IP ou au niveau applicatif. Le mécanisme proposé permet de mieux estimer les conditions réseaux pour appliquer une fragmentation adaptative minimisant ainsi le taux de perte globale. Ce mécanisme introduit aussi une limite inférieure de la taille d'un fragment MAC pour minimiser les pertes et optimiser l'utilisation de la bande passante.
- Garantie du débit d'émission vidéo par groupement de trames MAC sur la base de la structure d'image. Ce mécanisme exploite le TXOP (Transmission Opportunity) pour garantir plus de bande passante au niveau du lien d'accès 802.11 pour les flux vidéo. La transmission des trames MAC contenant des paquets vidéo se fait en rafale (burst), ce qui augmente le débit d'émission des flux vidéo comparé à d'autres flux concurrents.

4.3.1 Adaptation du débit vidéo en fonction du débit physique 802.11

Dans un environnement 802.11 où les utilisateurs sont mobiles, le débit physique du lien peut varier considérablement. Ce débit peut atteindre 1 Mbps dans certaines situations. Nous avons effectué des expérimentations sur le partage du débit entre plusieurs stations dans un environnement 802.11 et nous avons constaté que la station disposant du plus faible débit tire les autres stations vers un débit plus faible. Prenons l'exemple de trois stations utilisant trois débits différents. Station 1 avec un débit de 1 Mbps, station 2 avec un débit de 11 Mbps, et station 3 avec un débit de 54 Mbps. Le débit effectif réel utilisé pour chaque station est de l'ordre de 800Kbps. La première raison qui explique ce débit faible est due au partage équitable du canal de transmission entre toutes les stations. La même quantité de données doit être envoyée pour chaque station durant un intervalle de temps fixe. La deuxième raison est liée au débit physique lui-même puisque la station qui

transmet avec un débit faible (station 1 dans notre exemple) occupe le canal pour une durée plus longue. Ce problème introduit de nouvelles exigences pour maintenir la QoS des flux vidéo. D'une part, la nécessité d'informer la couche applicative sur la variation du débit physique des stations. D'autre part, le besoin d'estimer dynamiquement le débit physique effectif de chaque station afin d'adapter efficacement les flux qui leurs sont destinés. Enfin, la nécessité de mettre en œuvre un mécanisme qui permet d'adapter le débit vidéo en temps réel. Ceci est l'objet de notre première contribution.

Ainsi, nous avons mené plusieurs expérimentations pour essayer de comprendre la notion de partage de la bande passante dans les réseaux 802.11. À partir de ces études, nous avons proposé un modèle mathématique pour estimer le débit physique effectif pour chaque station au niveau d'un point d'accès (AP) 802.11.

Comme indiqué précédemment, le partage de la bande passante doit être équitable entre toutes les stations utilisant le point d'accès. Ainsi, à une charge du réseau maximale et pour la même taille de trame, chaque station doit transmettre exactement le même nombre de trames sur une durée de temps fixe, soit 1 seconde. L'équation 4-1 illustre le partage équitable de la bande passante sur un slot de temps d'une seconde.

$$\frac{Nbf \ fs}{r_1} + \frac{Nbf \ fs}{r_2} + \frac{Nbf \ fs}{r_3} + \dots + \frac{Nbf \ fs}{r_i} + \dots + \frac{Nbf \ fs}{r_{Nb_{sta}}} = 1s \quad \text{Eq. 4-1}$$

- Nbf est le nombre de trames transmises,

- fs est la taille d'une trame,

- r_i : Le débit physique théorique utilisé pour transmettre les trames à la station i .

Le débit physique effectif réalisé par chaque station Ri_{phy} est calculé par l'équation 4-2.

$$Ri_{phy} = Nbf \ fs / 1s \quad \text{Eq. 4-2}$$

D'où

$$Nbf = \frac{1}{fs \sum_{i=1}^{Nb_{sta}} \frac{1}{r_i}} \quad \text{Eq. 4-3}$$

En remplaçant l'expression du Nbf dans l'équation Eq. 4-2, le calcul du Ri_{phy} peut être simplifié en utilisant l'équation Eq. 4-4.

$$Ri_{phy} = \frac{1}{\sum_{i=1}^{Nb_{sta}} \frac{1}{r_i}} \quad \text{Eq. 4-4}$$

Le débit physique doit être supérieur au débit applicatif puisque les en-têtes d'encapsulations ne sont pas considérés comme l'*overhead* introduit par le mécanisme d'accès au canal CSMA/CA. L'équation 4-4 permet une bonne estimation du débit physique avec une erreur acceptable. Si on applique cette équation sur l'exemple de trois stations utilisant trois débits différents. À savoir : station 1 avec un débit de 1 Mbps, station 2 avec un débit

de 11 Mbps, et la station 3 avec un débit de 54 Mbps, nous trouvons un débit de 924Kbps comparé à celui trouvé par expérimentation qui est de l'ordre de 800Kbps.

Cette estimation de débit a été implémentée dans notre système **MIM (Meet In the Middle)** pour permettre à la source vidéo de s'adapter à la variation de la bande passante disponible. Elle est testée sur trois types d'algorithmes de contrôle de débit RCA : *Onoe*, *Amrr* et *Sample*. Les résultats montrent que notre modèle d'estimation du débit minimise les pertes correctement avec les algorithmes *Onoe* et *Amrr*, vu leurs stabilités dans le basculement d'un débit vers un autre, alors qu'il est moins performant avec l'algorithme *Sample* qui fait un choix aléatoire dans le changement de débit physique.

4.3.2 Adaptation conjointe FEC/débit vidéo en fonction de la puissance du signal et des taux de pertes

Nous avons évoqué précédemment que la fiabilité des communications sans fil représente un des principaux obstacles pour la distribution des flux vidéo avec QoS. Ainsi, la technique de correction des erreurs à la réception appelée FEC (Forward Error Correction) représente un mécanisme parfait pour fiabiliser les flux vidéo. La technique FEC est basée sur l'utilisation d'un code correcteur d'erreurs. Le flux initial est codé, puis transmis sur le canal et finalement décodé à l'arrivée.

Pour la transmission des données, les codes correcteurs d'erreurs peuvent être appliqués sur deux niveaux distincts : au niveau bit et/ou au niveau paquet. Dans le premier niveau, le symbole utilisé dans le codage est un bit et le code correcteur génère des bits redondants. Par contre, dans le deuxième niveau, un paquet est considéré comme un symbole et des paquets redondants sont codés à partir des paquets originaux.

La FEC au niveau paquet, plus précisément lorsqu'elle est appliquée au niveau applicatif/transport (exemple au niveau RTP), est plus avantageuse pour le transport des flux vidéo sur les réseaux sans fil pour les différentes raisons suivantes :

- L'unité de donnée de base dans un réseau est un paquet. Deux situations existent : soit un paquet est correctement reçu, soit il présente des erreurs et sera considéré comme perdu et rejeté lors de la vérification du CRC. Ainsi, les pertes dans le réseau s'effectuent par paquet et la FEC au niveau bit n'est pas adaptée pour décoder un paquet perdu entièrement.
- Lorsque la FEC au niveau bit est utilisée, les paquets corrompus durant la transmission, principalement dans les réseaux sans fil, n'atteignent jamais la couche applicative puisqu'ils sont supprimés automatiquement au niveau MAC durant la vérification du CRC.
- La FEC au niveau paquet offre une capacité de décodage supérieure à celle d'une FEC au niveau bit. Un paquet redondant, codé à partir d'un groupe de paquets, peut décoder n'importe quel paquet perdu dans ce groupe, ce qui n'est pas le cas pour une FEC au niveau bit.

Les inconvénients engendrés par l'utilisation de la FEC résident dans l'augmentation du débit de la source (débit du flux vidéo) et l'augmentation du délai d'émission des paquets. L'augmentation du débit de la source est due principalement au taux de redondance utilisé au niveau du processus de codage de la FEC, tandis que l'augmentation du délai de transmission est due à l'introduction de délais supplémentaires pour regrouper et **coder** les paquets ainsi que pour regrouper et **décoder** les paquets.

Il est clair que l'utilité de la FEC se justifie en présence de pertes dans le réseau. Cependant l'augmentation du débit de la source peut, à son tour, provoquer plus de pertes. Cette situation peut entraîner un effet boule de neige où les paquets FEC provoquent, à leur tour, des pertes de données. En revanche, l'utilité de la FEC est réduite lorsque l'état du canal s'améliore. Il faut noter que dans notre système, les pertes enregistrées au niveau récepteur sont signalées à la source vidéo périodiquement par l'utilisation du protocole RTCP, plus exactement dans les rapports RR (Receiver Report).

C'est pour ces raisons que nous avons conçu un système adaptatif qui se base sur plusieurs indicateurs pour optimiser la fiabilité du flux vidéo. Ainsi, notre contribution préconise un ajustement dynamique du taux de la redondance FEC toute en gardant un débit stable au niveau de la source vidéo. Ce mécanisme d'adaptation conjointe FEC/débit vidéo agit sur la source vidéo en réduisant son débit par un mécanisme de « transrating » et compense les pertes en ajoutant de la FEC. Le système redevient à son état initial si le réseau redevient à nouveau fiable. Ce mécanisme est une extension significative de la contribution précédente, dans laquelle la source vidéo est consciente de son environnement de transmission sous-jacent. Pour cela, nous avons modélisé le canal physique par deux états : un état bon (Good) et un état mauvais (Bad). Les deux états sont déterminés en fonction de la puissance du signal reçu. Nous avons utilisé le SSI (Signal Strength Indicator) comme indicateur de l'état du canal. Dans chaque état, la source vidéo ajuste son débit et son taux de redondance FEC.

Pour illustrer ce principe d'adaptation conjointe FEC/débit, la Figure 4.1 montre trois situations d'adaptation : (1) canal dans un état « bon », (2) canal dans un état « mauvais », et (3) à nouveau le canal dans un état « bon ». Regardons la situation (2) dans laquelle le canal est dans un état « mauvais », dans ce cas, l'ajout de la FEC ne provoque pas une augmentation du débit à la sortie de la source vidéo. Ceci est intéressant pour maintenir le débit initial de la source et réduire, de ce fait, l'augmentation du débit due à l'ajout de la FEC et donc les pertes.

Erreur ! Des objets ne peuvent pas être créés à partir des codes de champs de mise en forme.

Figure 4.1 : L'adaptation conjointe du taux de redondance FEC et du débit vidéo

La Figure 4.2 illustre le déroulement du processus d'adaptation conjointe FEC/débit pour fiabiliser le transport de la vidéo. Dans chaque état : « GOOD » ou « BAD », des adaptations FEC/débit sont appliquées. En effet, suivant les taux de perte rapportés par RTCP, le taux de redondance FEC est ajusté et agit ainsi sur le débit vidéo.

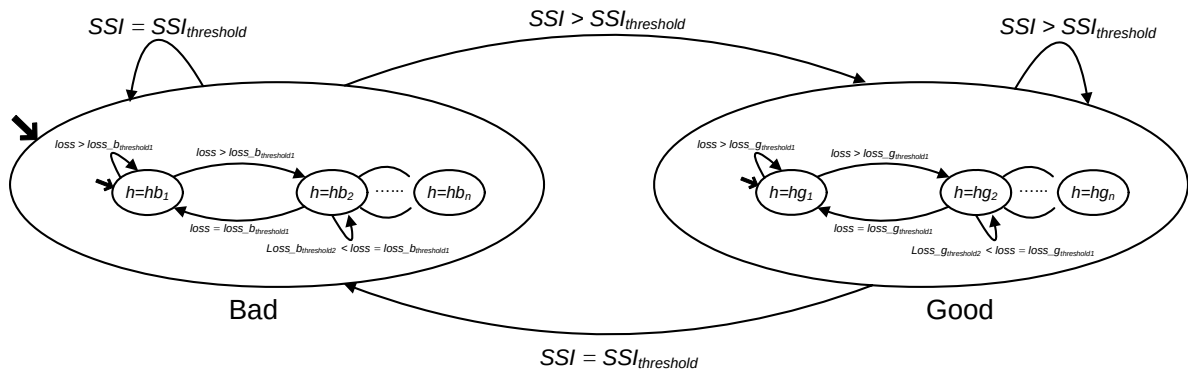


Figure 4.2 : Les différents états du système FEC adaptatif

Les expérimentations d'adaptation conjointe FEC/débit montrent qu'en diminuant le débit de la source vidéo, nous diminuons aussi la qualité perçue. En revanche le débit global du flux vidéo au niveau de la source reste stable, mais il devient plus résistant face aux pertes qui dégradent la qualité vidéo.

4.3.3 Fragmentation adaptative cross-layer au niveau MAC 802.11

Les erreurs binaires sur une trame MAC 802.11, qui sont dues aux bruits et aux interférences dans le canal, dégradent considérablement la qualité visuelle du flux vidéo. Les erreurs binaires sont détectées par la somme de contrôle (i.e. CRC) présente dans chaque trame. Une trame qui perd son intégrité est rejetée.

Il est clair que pour un taux d'erreur bit (BER) fixe, plus la trame est longue, plus la probabilité du rejet de celle-ci est importante. Ainsi, dans un environnement caractérisé par des erreurs binaires, la diminution de la taille de la trame à transmettre, par des mécanismes de fragmentation, est primordiale. En effet, la transmission d'une trame longue nécessite plus de temps, comparée à une trame plus petite. Une trame longue est plus exposée aux interférences et au bruit, en conséquence, elle présente une probabilité de corruption plus importante. Réduire la taille de la trame fait diminuer son temps de transmission sur le canal et par la même sa probabilité de corruption.

Cependant, deux inconvénients majeurs sont engendrés par le processus de réduction des tailles de trames à transmettre sur le canal 802.11. d'un part, l'augmentation du débit due aux différents en-têtes ajoutés (i.e. augmentation de l'overhead), et d'autre part, l'augmentation du délai de transit dans le réseau due au remplissage des différentes files d'attente traversées. D'un autre côté, la fragmentation est utile uniquement lorsque le système enregistre des pertes qui sont provoquées par les interférences, mais la fragmentation est contraignante lorsqu'il n'y a pas de pertes puisqu'elle introduit les inconvénients cités. C'est pour ces raisons que le contrôle du processus de fragmentation doit être dynamique et doit tenir compte de l'état du canal. Reste à déterminer la meilleure couche pouvant appliquer cette fragmentation.

Pour cela, nous avons mené une étude comparative pour déterminer la meilleure couche pouvant appliquer la fragmentation d'un paquet vidéo engendré par un codec vidéo. Trois niveaux sont candidats pour appliquer cette fragmentation :

- Niveau système ou applicatif, qui est pris en charge par le protocole ALF (Application Level Framing) qui permet d'ajuster la taille de la trame système à mettre dans un paquet RTP.
- Niveau réseau ou IP qui permet de découper suivant le MTU du réseau un segment provenant de la couche transport, dans notre cas, un datagramme UDP. La fragmentation duplique l'en-tête IP sur chaque fragment.
- Niveau support ou MAC qui utilise un mécanisme de fragmentation/réassemblage au niveau MAC qui permet de réduire la taille des trames transmises par la couche physique 802.11. Le standard définit le « *fragmentationThreshold* » qui représente la taille maximale d'une trame à transmettre.

Les résultats de notre étude démontrent que la fragmentation au niveau MAC présente de meilleures performances, comparé aux autres mécanismes de fragmentation. En effet, les pertes s'annulent complètement dans la fragmentation MAC puisque cette dernière se base sur un mécanisme d'émission des fragments en rafale appelé « Fragment Burst » qui transmet tous les fragments engendrés en une seule fois sans relâcher le contrôle du canal de transmission. Ceci réduit considérablement le temps d'attente des trames MAC dans la file d'attente, ce qui évite, par la même, sa surcharge et la suppression de trames en attente.

Une fois le niveau de fragmentation déterminé, nous nous sommes intéressés au calcul de la taille optimale d'un fragment MAC. Cette taille représente un compromis entre la nécessité de réduire les taux de perte et l'augmentation du débit due à « l'overhead » introduit par le mécanisme de fragmentation. En effet, les erreurs binaires (BER) d'un canal de transmission dépendent de plusieurs paramètres : la modulation du signal, la puissance du signal et le rapport de la puissance du signal sur bruit. Il est important aussi de signaler que chaque station dans le réseau expérimente un BER indépendant des autres stations et qui dépend principalement de l'état du canal lors de la transmission. Dans la pratique, il est impossible de déterminer la valeur du BER dans un environnement donné pour déterminer la taille de la trame optimale qui minimise les pertes. Cependant, nous avons développé un modèle mathématique permettant de déterminer la limite inférieure de la taille de la trame afin de minimiser les pertes et l'overhead. Nous avons vu dans la section 4.3.1, que le débit partagé entre stations se calcule en fonction des débits physiques des différentes stations connectées. Le modèle qui a été défini est le suivant :

$$Ri_{phy} = \frac{1}{\sum_{i=1}^{Nb_{sta}} \frac{1}{r_i}} \quad \text{Eq. 4-5}$$

Le taux d'overhead appelé Ri_{ovh} maximum peut être exploité par une station et représente le reste du rapport entre le débit du flux destiné à cette station, au niveau de la couche LLC (Logical Link Control) Ri_{llc} , et le débit physique Ri_{phy} utilisé par la station. Ceci est illustré par l'équation Eq. 4-6 :

$$Ri_{ovh} = 1 - \frac{Ri_{llc}}{Ri_{phy}} \quad \text{Eq. 4-6}$$

Une nouvelle équation Eq. 4-7 peut être obtenue à partir des équations Eq. 4-5 et Eq. 4-6.

$$Ri_{ovh} = 1 - Ri_{llc} \sum_{i=1}^{Nb_{sq}} \frac{1}{r_i} \quad \text{Eq. 4-7}$$

Si on considère que l'overhead d'une station peut être calculé à partir des tailles des en-têtes au niveau MAC (hs_{mac}) et PHY (hs_{phy}) en fonction de la taille d'une trame (fsi) nous obtenons cette équation :

$$Ri_{ovh} = \frac{hs_{mac} + hs_{phy}}{fsi} \quad \text{Eq. 4-8}$$

Ainsi, nous pouvons en déduire la limite inférieure de la taille d'une trame à partir des équations Eq. 4-8 et Eq. 4-7

$$fsi_{threshold} = \frac{hs_{mac} + hs_{phy}}{1 - Ri_{llc} \sum_{i=1}^{Nb_{sq}} \frac{1}{r_i}} \quad \text{Eq. 4-9}$$

L'équation Eq. 4-9, donne la limite inférieure d'une taille de trame qui à la fois minimise l'overhead et ne casse pas la relation de partage du débit équitable entre stations. Les résultats d'expérimentation montrent la capacité de ce mécanisme de fragmentation adaptative à diminuer les pertes contre une augmentation très limité du débit d'émission du flux vidéo. C'est ainsi que la qualité visuelle du flux vidéo est maintenue lorsque le canal subit des interférences.

4.3.4 Groupage de trames basé sur une image vidéo

Les flux vidéo nécessitent beaucoup de ressources et en particulier un débit relativement élevé. Dans un environnement à ressources partagées et fluctuantes, tel que l'environnement 802.11, l'adaptation du débit vidéo pendant la transmission est importante. Cette adaptation a été présentée dans les contributions précédentes. Toutefois, les mécanismes présentés, précédemment, ne permettent pas de garantir au flux vidéo, un débit plus élevé au niveau du lien d'accès. Dans le cas où le flux vidéo est concurrencé par d'autre flux UDP ou TCP, le débit du flux vidéo au niveau du lien d'accès 802.11 sera drastiquement limité à cause du partage équitable du lien d'accès garanti par le mécanisme d'accès DCF (Distributed Coordination Function) qui utilise la méthode d'accès CSMA/CA. Dans cette contribution, nous proposons un mécanisme assurant une certaine correspondance entre le débit vidéo et le nombre de trames qui sont envoyées durant un intervalle d'opportunité

de transmission TXOP. Ce mécanisme permet aux flux vidéo d'avoir des (TXOP) adaptatives en fonction de leur débit grâce à des transmissions en rafales (burst) des trames MAC appartenant à la même image vidéo. Ce mécanisme de groupage de trames consiste à rassembler, dans un seul envoi au niveau MAC, les trames qui possèdent le même destinataire en partageant le temps d'accès au canal. Ceci permet aux trames d'être transmises en rafale sans relâchement du canal de transmission et chaque trame est acquittée séparément de la part du récepteur.

Cependant, cette technique n'est pas adaptée dans un environnement qui présente des erreurs binaires et elle va à l'encontre du mécanisme de fragmentation présenté dans la contribution précédente. Ainsi, le nombre de trames à concaténer doit être contrôlé et doit dépendre de l'état du canal. La Figure 4.3 illustre le principe de fonctionnement du mécanisme de groupage de trames, comparé à un processus d'envoi classique. L'ensemble des trames appartenant au même groupe bénéficie d'un envoi express de la même manière que dans le cadre de la fragmentation MAC. Notre approche constitue un groupe de trames appartenant à la même image vidéo.

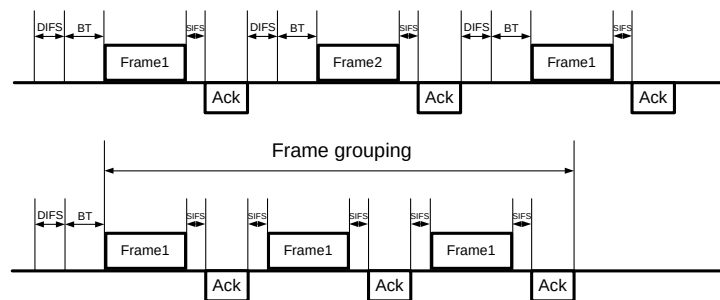


Figure 4.3 : Principe de groupage de trames

Ce mécanisme présente deux défis majeurs pour déterminer la taille d'un groupe de trames et délimiter ainsi l'intervalle de temps durant lequel les trames sont transmises en rafales. Le premier défi est la prise en charge du débit réel du flux vidéo. Cette prise en charge est assurée, dans notre mécanisme, par le groupage des trames sur la base de la structure en image. De plus, nous assurons une priorité d'occupation du lien d'accès en relation avec la hiérarchie temporelle des images I, P, et B. Le deuxième défi consiste à éviter la monopolisation du canal de transmission lors du groupage de trames. Ceci ne peut pas se produire dans notre cas puisque le nombre de paquets dans une image est limité. La Figure 4.4 illustre un exemple de groupage de trames à base d'une image vidéo.

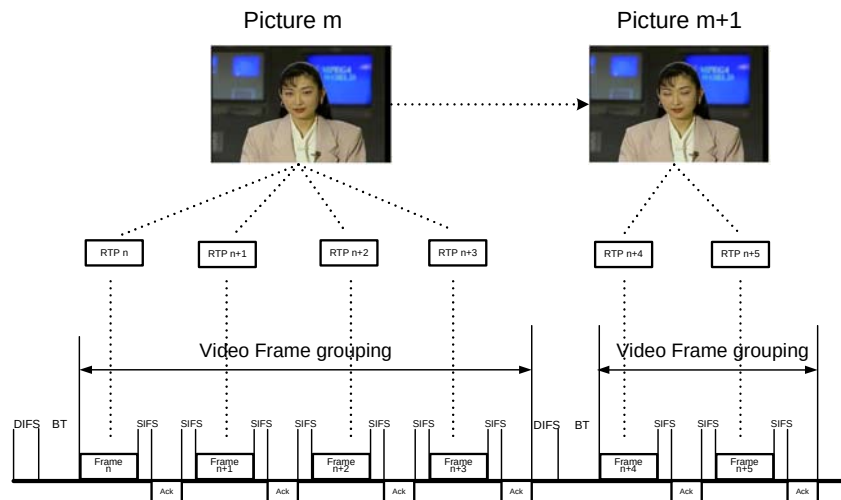


Figure 4.4 : Groupage de trames MAC à base d'images vidéo

Il est important de noter que la durée du TXOP, dans notre cas, n'est pas fixe mais dépend de la taille du groupe de trames, alors que le mécanisme 802.11e introduit des catégories d'accès avec des TXOP fixes ; là est la différence majeure. Dans le mécanisme 802.11e, la variation du débit physique n'influence pas la durée du TXOP, ceci peut s'avérer contraignant si le débit physique chute.

Le groupage de trames à base d'images exploite des informations « in-band » pour déterminer les trames appartenant à la même session et à la même image vidéo. Les évaluations de ce mécanisme ont montré : (1) l'avantage d'avoir un mécanisme de groupage adaptatif en relation avec le débit de la vidéo, et (2) la capacité de ce mécanisme à occuper le canal pour satisfaire le débit vidéo. Les résultats montrent, clairement, que le groupage de trames favorise les flux vidéo et leur donne plus de priorité par rapport à d'autres flux. Les gains obtenus sont la réduction des taux de pertes et la garantie d'un débit élevé au niveau du lien d'accès.

4.4 Conclusion

Les réseaux locaux sans-fil 802.11 se généralisent, s'imposent dans notre vie quotidienne et attirent de plus en plus les services multimédia temps réel (TV, VoD, Téléphonie, etc.). Malgré l'accroissement de la bande passante, et la perspective d'émergence prochaine du nouveau standard 802.11n qui permettra, à terme, des débits atteignant 540 Mbps (débit physique), le transport des services vidéo temps réel sur ce type de réseaux présente des défis importants. Les effets négatifs des erreurs et pertes de données et la variation des ressources disponibles dans le canal physique influencent considérablement la qualité de service. C'est pourquoi nous avons conçu un système adaptatif de transport de flux vidéo temps réel sur les réseaux 802.11 basé sur l'approche cross-layer. Ce système appelé MIM (Meet In the Middle) intègre des mécanismes avancés d'adaptation de traitement de transport vidéo sur les couches protocolaires importantes. Quatre mécanismes importants ont été

implémentés et validés dans notre système : (1) le mécanisme d'adaptation du débit du service vidéo pour refléter la variation du débit physique, (2) le mécanisme d'adaptation conjointe FEC/débit permettant de remédier aux problèmes d'erreurs et de pertes de paquet, (3) le mécanisme de fragmentation adaptative au niveau MAC 802.11 pour remédier au problème des erreurs dans le canal, et (4) le mécanisme de groupement de trames MAC pour garantir un débit important au flux vidéo.

4.5 Mes Publications Associées

- [1] Ismail Djama, *Toufik Ahmed*, A. Nafaa and R. Boutaba, "Meet In the Middle Cross-Layer Adaptation for Audiovisual Content Delivery", in the **IEEE Transactions on Multimedia**, Volume 10, Issue 1, pp. 105 – 120, Jan. 2008.
- [2] Ismail Djama and *Toufik Ahmed* "A Cross-Layer Interworking of DVB-T and WLAN for Mobile IPTV Service Delivery" to be published in **IEEE Transaction on Broadcasting**, ISSN 0018-9316, Vol. 53 Issue: 1, Part: 2, pp. 382-390, march 2007.
- [3] Ismail Djama, *Toufik Ahmed* "MAC-Level Video Frame Grouping using Cross-Layer Architecture" to appear in IEEE International Symposium on Wireless Pervasive Computing (**ISWPC**), Santorini Greece, May 2008.
- [4] Ismail Djama and *Toufik Ahmed* "Adaptive Cross-layer Fragmentation for Reliable Wireless IPTV Services" in proce. of IEEE International Global Information Infrastructure Symposium (**GIIS**), pp. 203-206, Marrakech, 2-6 July 2007.
- [5] Ismail Djama and *Toufik Ahmed* "MPEG-21 Cross-Layer QoS Adaptation for Mobile IPTV Service Delivering" in 2nd International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution (**AXMEDIS'06**), 12-15 Dec. 2006. URL : <http://www.axmedis.org/axmedis2006/docs/AXMEDIS2006-programme-ver2-8-FINAL.pdf>
- [6] *Toufik Ahmed*, Ismail Djama "Delivering Audiovisual Content with MPEG-21-Enabled Cross-Layer QoS Adaptation" in Packet Video 2006 published in **Journal of Zhejiang Univ SCIENCE A** 7(5) from **IEEE Packet Video PV**, pp. 784 784-793. April 2006.
- [7] Ismail Djama, *Toufik Ahmed*, Daniel Negru, Nawel Zangar "An MPEG-21-enabled Video Adaptation Engine for Universal IPTV Access" in **IEEE International Symposium on Broadband Multimedia Systems and Broadcasting 2006**, Las Vegas, NV USA April 6-7, 2006. URL: <http://www.ieee.org/organizations/society/bt/06bmsympo/program.html>

4.6 Bibliographie

- [SRI 05] V. Srivastava and M. Motani, "Cross-layer design: a survey and the road ahead", IEEE Communications Magazine, vol.43, no.12, pp.112-119, December 2005.
- [VAN 05] M. Van Der Schaar, et al., "Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms", IEEE Wireless Communications Magazine, vol. 12, no. 4, pp. 50-58, August 2005.
- [SED 06] Nasser Sedaghati-Mokhtari, Mahdi Nazm Bojnordi, Nasser Yazdani, "Cross-Layer Design:A New Paradigm", Proc In International Symposium on Communications and Information Technologies (ISCIT '06), On page(s): 183-188, Bangkok, Sept 2006.

- [WAN 03] Q. Wang, M. A. Abu-Rgheff, "Cross-Layer Signalling for Next-Generation Wireless Systems," Proc In IEEE Wireless Communication and Networking , New Orleans, Mar. 2003.
- [WIN 06] R.Winter, J.H.Schiller, N.Nikaein, C.Bonnet, "CrossTalk: cross-layer decision support based on global knowledge", in IEEE Communications Magazine, ISSN: 0163-6804, pages 93- 99, Volume: 44, Issue: 1, Jan 2006.
- [KHA 06] S.Khan, Y.Peng, E.Steinbach, M.Sgroi, W.Kellerer, "Application-driven cross-layer optimization for video streaming over wireless networks", In IEEE Communications Magazine, ISSN: 0163-6804, pages 122-130, Volume: 44, Issue: 1, Jan 2006.
- [4MORE] 4MORE project website <http://4more.av.it.pt/>
- [PHO] PHOENIX project website <http://www.ist-phoenix.org/>
- [ENT] ENTHRONED II project website <http://www.ist-enthrone.org/>
- [KAM 97] Kamerman and L. Monteban. WaveLAN-II: "A high-performance wireless LAN for the unlicensed band", AT&T Bell Laboratories Technical Journal, pages 118-133, 1997.
- [LEC 04] M. Lacage, M. H. Manshaei, and T. Turletti, "IEEE 802.11 rate adaptation: a practical approach," in Proceedings of the 7th Symposium on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM '04), pp. 126–134, Venice, Italy, October 2004.
- [PAV 03] Jd.P.Pavon, Sunghyun Choi, "Link adaptation strategy for IEEE 802.11 WLAN via received signal strength measurement", IEEE International Conference on Communications (ICC '03), Volume: 2, On page(s): 1108-1113, Anchorage, AK, USA, 2003.
- [HOL 01] G. Holland, N. Vaidya and P. Bahl, "A Rate-Adaptive MAC Protocol for Multi-Hop Wireless Networks", in Proceedings of the 7th annual international conference on Mobile computing and networking (ACM MOBICOM'07), Pages: 236 - 251, Rome, July 2001.
- [HAR] Haratcherev, J. Taal, K. Langendoen, R. Lagendijk and H. Sips, "Automatic IEEE 802.11 rate control for streaming applications", Wireless Communications and Mobile Computing, Vol 5, pp.412-437, 2005.

5 Transport et Contrôle de la Qualité des Services Vidéo sur les réseaux P2P

5.1 Introduction

Le réseau Internet se développe, change, s'organise et innove constamment de nouvelles applications et de nouveaux services. Ces dernières années, l'un des événements qui a le plus marqué la croissance d'Internet est certainement le développement des réseaux Pair-à-Pair ou Peer-to-Peer (P2P). Dans ce genre de réseaux, les utilisateurs s'organisent en groupe, en communauté ou en tant qu'utilisateurs simples. Le concept de cette organisation repose sur le fait que chaque participant au réseau P2P est un consommateur (client) et un fournisseur potentiel (serveur). Les utilisateurs ne sont plus passifs, consommant simplement des flux de données en provenance de quelques serveurs répartis sur Internet, mais ils sont actifs dans la mesure où ils participent à la création, la génération et à la diffusion de contenus.

Cette croissance grandissante des réseaux P2P suscite, actuellement, de plus en plus d'attention de la part de la communauté des chercheurs et s'est traduite par l'apparition d'un nombre important de travaux de recherche, de projets et de groupes de travail. Les réseaux P2P se présentent comme une réelle alternative au traditionnel modèle client / serveur, au vu de leur capacité à assurer la montée en charge (scalability), et leur déploiement rapide à faible coût.

Ainsi, dans notre troisième volet de recherche, nous nous sommes intéressés à la distribution de flux vidéo temps réel sur les réseaux P2P. Ce travail a été mené dans le cadre de la thèse de doctorat de M. Mubashar Mushtaq. Plusieurs mécanismes ont été proposés pour le contrôle de la QoS dans le transport multi-sources des flux vidéo sur les réseaux P2P, à savoir :

- Un mécanisme de sélection des meilleurs pairs (nœuds) participant à la session de diffusion vidéo. Ceci permet la réception du flux à partir des pairs disposants de la meilleure QoS [35]. Ce mécanisme se base sur des mesures actives permanentes de la qualité de bout-en-bout des liens entre le récepteur et les clients pairs potentiels. Les pairs disposant de la même qualité de service sont organisés ensemble dans un même réseau overlay pour faciliter leur recherche et leur sélection. Ainsi, ce mécanisme est accompagné de stratégies de regroupement de pairs en fonction de différents critères (sémantique et physique) afin d'améliorer le fonctionnement global du système de diffusion vidéo.

- Un mécanisme de commutation de flux et de pairs qui permet de maintenir un niveau de QoS acceptable en cas de problème de réception dû au vidage imprévu des mémoires tampons de réception [2]. Dans ce cas, deux solutions se présentent : (1) dégrader la qualité du flux reçu, par des mécanismes d'adaptation vidéo dans le cadre du codage hiérarchique SVC ou (2) essayer de maintenir la qualité originale par l'identification des pairs en défaillance et le remplacement par des pairs assurant la qualité originale.
- Un mécanisme de réservation de ressources de bout-en-bout assisté par le fournisseur de services. Ce mécanisme permet au fournisseur de services (SP) et au réseau P2P de collaborer pour mieux servir l'utilisateur final.

5.2 État de l'art

Ces dernières années, les réseaux P2P se sont généralisés. Ils se sont présentés comme une formidable opportunité de croissance du marché de l'accès haut-débit à Internet. Ils offrent aussi des possibilités et des perspectives de développement de nouvelles applications. Grâce aux réseaux P2P, de plus en plus de personnes partagent leurs documents, photos, vidéos, musiques, et autres données via Internet. Plusieurs systèmes ont ainsi vu le jour pour supporter cette nouvelle donne. Le phénomène P2P a vraiment révolutionné la façon dont les utilisateurs consomment et interagissent avec l'information en particulier les flux multimédia. Actuellement, on estime que le trafic Internet est dominé à 70% par des applications P2P. Cette proportion de trafic est destinée à évoluer dans le futur vu la croissance continue du nombre d'utilisateurs connectés à Internet via divers types de réseaux d'accès (ADSL, WiFi, optique, LAN, etc.), et qui utilisent différents types de terminaux d'accès (PC, portable, téléphone, etc.). Cette hétérogénéité dans la technologie et les moyens d'accès à l'information pose un certain nombre de problèmes relatifs au contrôle de QoS et, en particulier, à l'adaptation des applications.

Afin de mieux appréhender les verrous spécifiques à l'utilisation des réseaux P2P dans le cadre du transport des flux vidéo sur Internet, cette section a pour but de dresser un état de l'art permettant de comprendre cette architecture de réseaux.

5.2.1 Taxonomie des réseaux P2P

Les réseaux P2P présentent trois caractéristiques importantes : auto-organisation, communication symétrique, et contrôle distribué [ROU 04]. L'auto-organisation des pairs (ou nœuds) est assurée par un processus de découverte qui permet le regroupement logique des nœuds. La communication symétrique est assurée par les nœuds qui jouent un rôle identique dans l'offre et la demande de services. C'est la différence majeure avec le modèle client / serveur dans lequel le client et le serveur ont des rôles différents. Dans le contrôle distribué, il

n'existe pas de contrôleur central qui dicte le comportement de certains nœuds. Ainsi, chaque nœud détermine son niveau de participation au réseau selon ses besoins. Un nœud décide, lui-même, quand accéder au réseau, quand le quitter et quels ressources partager et consommer. Ceci nécessite une distribution du contrôle entre tous les nœuds.

En plus de ces caractéristiques, les réseaux P2P peuvent être classés selon plusieurs critères et métriques, à savoir partiellement centralisé, décentralisé, et hybride. La catégorie décentralisée peut encore être divisée en décentralisé mais structuré, décentralisé et non structuré, ou décentralisé mais semi-structuré. La différence principale entre ces systèmes réside dans le mécanisme utilisé pour rechercher des ressources dans le réseau P2P.

Selon le degré de décentralisation, on peut classer les systèmes P2P comme suit :

- Système centralisé : dans ce système, la description des ressources partagées, appelées aussi index de recherche, est disponible dans un annuaire centralisé. Pour localiser une ressource, un nœud demande au serveur central la disponibilité de la ressource et son adresse éventuelle. Une fois localisée, la consommation de la ressource peut commencer directement à partir des pairs contenant celle-ci sans passer par l'annuaire centralisé. L'inconvénient de ce genre de système est le risque de surcharge de l'annuaire qui peut devenir un point de défaillance du système.
- Système décentralisé : dans ce système, il n'existe pas d'annuaire central. La recherche d'une ressource doit être envoyée aux nœuds voisins, qui eux même la font suivre à leurs voisins. On parle d'inondation du réseau jusqu'à ce que la ressource soit localisée. Il est clair que ce genre de système est très résistant à la défaillance des nœuds, mais il présente un inconvénient majeur relatif au temps de recherche important et au passage à l'échelle. Pour simplifier la recherche, les nœuds dans ce réseau peuvent être reliés logiquement entre eux pour constituer un réseau overlay qui permet de faciliter le processus de recherche. On parle alors d'un réseau décentralisé mais structuré. L'emplacement de la ressource peut être déterminé par des tables de hachage distribuées (ou DHT Distributed Hash Table) qui sont généralement calculées à partir des mots contenus dans la ressource partagée.
- Système Hybride : ce système combine l'efficacité du mécanisme de recherche des systèmes centralisés et le passage à l'échelle des systèmes décentralisés. Il est appelé partiellement centralisé au vu de l'utilisation de quelques pairs jouant le rôle de serveurs centraux et qui sont aussi appelés super-nœuds (super-node). Un super-nœud est généralement élu par ses pairs sur des critères de performance et de bande passante. Il doit connaître tous les nœuds qui lui sont logiquement rattachés et autres super-nœuds. Actuellement, les systèmes hybrides sont largement utilisés dans des produits tels que : KaZaA, BitTorrent, eDonkey/eMule.

5.2.2 Applications utilisant les réseaux P2P

De nombreuses applications utilisent les réseaux P2P pour faciliter la distribution du contenu vers un grand nombre d'utilisateurs. Nous les avons classées en plusieurs domaines d'application :

- Applications de distribution de contenus (content delivery application). Le but de ces applications est la distribution massive de contenu pour alléger la surcharge des serveurs prévus initialement à cette tâche. La plupart des produits P2P connus rentrent dans cette catégorie à savoir : FastTrack (KaZaA, Morpheus), Napster, eDonkey/eMule, Akamai, BitTorrent, etc. Le contenu multimédia (vidéo, images, audio) et les applications logicielles sont les deux ressources les plus partagées par ce genre d'applications. En ce qui concerne la distribution des flux multimédia, son principe de fonctionnement est simple. Le fichier doit être complètement téléchargé avant sa consommation par l'utilisateur. Ce qui pose certains problèmes relatifs au temps de téléchargement, à l'occupation de l'espace disque, à la qualité et au type de contenu. En revanche, actuellement, la notion de service de diffusion temps réel de contenu multimédia voit le jour avec l'apparition de plusieurs applications telles que : CoopNet, SpreadIt, ZIGZAG, PALS, and PROMISE. Le principe de fonctionnement repose sur les bases de diffusion temps réel de flux multimédia (i.e., téléchargement et lecture simultanée d'un flux). Dans le même contexte, certains produits se sont spécialisés dans la diffusion des flux TV sur les réseaux P2P tels que : Pplive, Zattoo, Gridmedia. **C'est dans le cadre de la diffusion de la vidéo pré-enregistrée que nous avons proposé des mécanismes de contrôle et d'adaptation de la QoS sur les réseaux P2P.**
- Applications de communication et de messagerie instantanée. Ces applications se développent largement sur Internet grâce aux facilités apportées par les réseaux P2P en terme de groupement d'utilisateurs, recherche distribuée, et efficacité de distribution. Le succès de Skype est un exemple concret de l'apport des réseaux P2P dans la gestion de la signalisation pour le service téléphonie sur IP entre les nœuds participants.
- Applications de calculs distribués. Le but de ces applications est de profiter de l'architecture P2P pour faire des calculs distribués. Chaque nœud connecté au réseau dispose d'une puissance de traitement disponible (cycle CPU) qui peut être exploitée pour faire des calculs complexes. L'objectif est de grouper tous les cycles CPU disponibles chez les utilisateurs (i.e., les pairs) pour en créer un supercalculateur capable de faire des traitements puissants. Plusieurs projets tels que Seti@Home [SET], Genome@Home [GEN] et Folding@Home [FOL] sont des exemples d'applications qui exploitent ce principe. Lors du développement de ce type de calcul distribué, une attention particulière doit être portée sur le problème d'exactitude des résultats. N'importe quel nœud peut prétendre fournir des résultats justes de calcul. Pour éviter ce problème, certaines solutions réalisent le calcul « n » fois sur des

nœuds différents. Si un certain nombre de résultats diffèrent, le calcul est considéré comme faux et il doit être refait sur d'autres nœuds de confiance.

- Applications de collaboration. La collaboration entre les utilisateurs pour accomplir un objectif commun constitue la base pour le développement de ce genre d'application. Ceci permet à plusieurs utilisateurs de travailler sur le même document en même temps grâce aux réseaux P2P sans passer par un serveur central. L'application Groove de Microsoft [GRO] est un exemple concret de ce type de solution.

Le Tableau 5.1 résume certains domaines d'application avec quelques exemples de produits.

Domaines d'applications	Exemples d'applications
Distribution de contenus	FastTrack (KaZaA, Morpheus), Napster, eDonkey/eMule, Akamai, BitTorrent
Diffusion multimédia	Joost, Narada, PPStream, PROMISE, PALS, GNUSstream
Services IPTV	TVUPlayer, Joost, CoolStreaming, Tribler, GridMedia, iGridMedia, Cybersky-TV, TVants, PPLive, LiveStation, Pando, Zattoo,
Calculs distribués	Seti@Home, Genome@Home, GridSystems, Avaki, Jivalti, Entropia, DataSynapse, Distributed.Net, DistributedScience, Popular Power, Ubero, 2AM
Applications de communication et de messagerie instantanée	Skype, AOL, AIM, ICQ, NetNews
Collaboration	Groove, Center Span, eZ, Interbind, Engenia Software, Inc. Jabber
Protocoles	JXTA, Jabber
Recherche distribuée, moteur de recherche	Next Page, Open Cola Folders, Pelbio, aKa InfraSearch, WebV2, sciencenet, AmbientDB, and Xpeer

Tableau 5.1 : Classification des réseaux P2P selon leur domaine d'application.

5.2.3 Les systèmes de streaming multimédia sur P2P

Les systèmes de streaming P2P permettent la diffusion de flux audiovisuel temps réel à grande échelle en utilisant les réseaux P2P. Les performances de communication de tels systèmes peuvent être considérablement améliorées par rapport au modèle client /serveur classique. La classification des modèles de diffusion de flux audiovisuels sur les réseaux P2P est étudiée dans ce qui suit :

- Diffusion basée sur une arborescence ou multicast applicatif : dans ce mode de diffusion de flux audiovisuels, les nœuds sont organisés hiérarchiquement pour former un arbre. La source initiale du flux est située à la racine de l'arbre. Le flux est envoyé depuis la racine vers l'ensemble des feuilles de l'arbre. Les nœuds intermédiaires jouent le rôle de relais pour servir plusieurs utilisateurs. Ce genre de système émule le fonctionnement du multicast dans les réseaux IP, mais il présente quelques

inconvenients comme la surcharge du nœud parent, la synchronisation des nœuds, l'hétérogénéité des nœuds, etc. Les systèmes de streaming les plus connus utilisant un arbre sont : SpreadIt [DES 01], PeerCast [ZHA 07A], NICE [BAN 02], ESM [CHU 02], Overcast [JAN 00]. Ces systèmes diffèrent les uns des autres dans les mécanismes de création de l'arbre, dans le maintien de l'arbre et dans la gestion des départs /arrivées des nœuds. La profondeur de l'arbre peut augmenter exponentiellement en fonction du nombre d'utilisateurs, ce qui peut poser certains problèmes de délai. Par conséquent, pour une meilleure exploitation de ce modèle, la profondeur de l'arbre et le nombre de niveaux qui séparent un utilisateur de son parent doivent être contrôlés. Un autre inconvénient est lié à la gestion des utilisateurs qui rentrent et quittent le réseau fréquemment puisqu'il faut maintenir la structure en arborescence. Ceci peut engendrer une signalisation importante dans le réseau.

- Diffusion multi-arborescences : ce mode d'organisation répond aux problèmes soulevés dans la diffusion basée sur une arborescence. L'objectif est de balancer la charge de la diffusion du flux sur plusieurs arbres. Pour cela, chaque arbre contient un sous flux du flux original. Un nœud doit être connecté à plusieurs arbres en même temps pour recevoir le flux initial. On parle du streaming multi-sources dans lequel le flux consommé provient de plusieurs sources différentes. Cependant, les différentes racines de chaque arbre présentent des points de défaillance du système. Des exemples d'implémentations de ce mode de streaming existent tels que : SplitStream [CAS 03], CoopNet [PAD 02] [PAD 03] et P2PCast [P2PCAST] [NIC 03].
- Diffusion maillée (ou mesh) : ce mode d'organisation est plus robuste face à la défaillance des nœuds parents, mais il est plus complexe à mettre en place puisque chaque nœud doit maintenir une relation avec plusieurs nœuds pour former un réseau maillé. Un nœud peut télécharger un contenu depuis n'importe quels autres nœuds et peut contribuer par sa bande passante sortante pour accomplir la tâche de streaming des autres nœuds. Les plus connus des systèmes de streaming qui implémentent ce mode de distribution de contenus sont : Chainsaw [PAI 05], PRIME [MAG], and DONet/CoolStreaming [ZHA 05].

La diffusion multi-arborescences réside entre la simplicité de l'approche utilisant une seule arborescence avec ses limites et la rigidité de l'approche utilisant une diffusion maillée (mesh) avec ses avantages. Notre approche de streaming des flux vidéo rentre dans le mode de diffusion multi-arborescences. Ce choix a été motivé par l'utilisation du format de codage vidéo (SVC) qui organise le flux vidéo en couches. Ainsi, dans notre approche chaque couche SVC est mappée sur un arbre de diffusion P2P.

5.2.4 Problèmes de streaming vidéo sur P2P

Le streaming des flux vidéo sur les réseaux P2P présente des problèmes spécifiques, comparé au streaming en mode client/serveur. En plus de tous les problèmes hérités du protocole IP : bande passante non garantie, présence de perte dans le réseau, et délai de transmission et gigue non bornés, les problèmes dans les réseaux P2P sont liés à l'organisation du réseau et aux comportements non déterministes des nœuds.

En effet, le streaming dans les réseaux P2P doit prendre en compte quelques contraintes liées à la présence de plusieurs nœuds dans le réseau, disposants du contenu désiré. La disponibilité de plusieurs copies du flux sur différents nœuds pose le problème de sélection des meilleurs nœuds pouvant assurer la qualité de service au récepteur, en particulier, pour maximiser la bande passante reçue. Il existe plusieurs métriques objectives ou subjectives qui nous aident à sélectionner les meilleurs nœuds, telles que, la mesure de la bande passante disponible entre les nœuds sources et le nœud récepteur, le délai de transmission possible, le délai aller-retour (RTT), le nombre de sauts, etc. Généralement des mécanismes de monitoring permettent d'aider à l'estimation de ces métriques. L'emplacement géographique des nœuds sources, par rapport au nœud récepteur peut être utilisé comme contraintes supplémentaires. Sans rappeler qu'il est plus avantageux, aussi bien pour le consommateur que pour les différents fournisseurs réseaux, d'établir la session de streaming entre les nœuds géographiquement proches. Ceci permet d'éviter de surcharger le réseau, et par la même, de maximiser le débit des nœuds. Plusieurs outils disponibles sur Internet remplissent cette tâche de localisation géographique des nœuds. La commande « Traceroute » et la base de données « WHOIS » peuvent être utilisés. D'autres projets, aussi existent dans ce contexte tels que : NetWorldMap [NWM] et VisualRoute [VIS] qui exploitent des techniques et des bases de données avancées pour permettre la localisation géographique des nœuds. Dans une situation d'équité entre nœuds, le débit maximal qu'un réseau P2P peut atteindre correspond à la capacité « upload » totale de tous les nœuds divisée par le nombre des pairs. Par conséquent, l'exploitation de la capacité « upload » d'un nœud dans le réseau P2P est un problème majeur. Le modèle de diffusion basé sur une arborescence ne peut être utilisé dans ce contexte, les nœuds intermédiaires ne peuvent pas servir le flux si leur capacité « upload » est inférieur au débit de la vidéo. Les approches existantes [VEN 06], [GAR 03], [CUI 04], [LI 04], et [SMA 07] divisent le flux vidéo en tranches (slices) et optimisent la diffusion vidéo en distribuant ces tranches sur plusieurs arbres pour utiliser autant que possible la capacité « upload » des nœuds. L'approche présentée dans [ZHA 07B] a démontrée qu'avec un mécanisme de type « pull » utilisant un ordonnanceur de paquets, la qualité de service pouvait être assurée à plusieurs utilisateurs, simultanément.

Le deuxième problème spécifique aux réseaux P2P, est celui lié au comportement dynamique des nœuds participants au réseau. Le plus important est le comportement connu sous le nom de « free-riding » qui consiste à quitter le réseau une fois le téléchargement terminé. Dans un réseau P2P, un nœud décide à lui seul quand rejoindre ou quitter un réseau P2P, il peut décider de ne plus partager la ressource demandée, ou aussi de limiter

la bande passante « upload » utilisée pour le streaming. Ce comportement imprévisible des nœuds pose un réel problème pour maintenir la qualité de service. Certains mécanismes incitent et motivent les utilisateurs à rester connectés au réseau P2P pour servir d'autres utilisateurs. Les mécanismes d'incitation peuvent être implémentés en conservant l'historique des contributions et en tenant compte de la quantité des ressources partagées. Néanmoins, le comportement dynamique et imprévisible des utilisateurs doit être géré et contrôlé afin de servir efficacement les clients. Des solutions d'adaptations doivent être utilisées pour, d'une part, adapter le flux vidéo aux ressources disponibles dans le réseau, mais aussi alléger le comportement dynamique des nœuds. Ces mécanismes d'adaptation peuvent être combinés intelligemment avec des flux vidéo codés hiérarchiquement (exemple SVC) pour garantir un streaming continu.

Le troisième problème concerne l'hétérogénéité des utilisateurs. Actuellement, de plus en plus d'utilisateurs se connectent aux réseaux P2P en utilisant des terminaux de plus en plus petits (PC, mini PC portable, téléphone, PDA, etc.). Ce qui impose de concevoir des systèmes de streaming P2P capables de tenir compte des caractéristiques des terminaux utilisés pour mieux servir les clients et mieux optimiser les ressources de transport. Là aussi, les flux vidéo codés hiérarchiquement (i.e., SVC) sont mieux adaptés.

Finalement, le streaming sur les réseaux P2P doit maintenir une relation synchrone entre les nœuds. Il est plus complexe que le téléchargement P2P puisqu'il impose des contraintes temps-réel. En effet, chaque nœud doit savoir exactement quel paquet vidéo (ou tranche de flux) envoyer et à quel moment. Le récepteur doit jouer le rôle d'un coordinateur en dictant à chaque nœud l'ordonnancement nécessaire. Ceci nécessite aussi l'implémentation de mécanismes de signalisation pour assurer ces tâches. Un ordonnancement efficace permet d'éviter des situations dans lesquels la mémoire tampon du récepteur peut déborder (*overflow*) ou se vider complètement (*underflow*).

5.3 Contributions

Nous avons développé des mécanismes permettant de remédier à quelques problèmes cités dans la section précédente pour permettre le contrôle de QoS des flux vidéo transportés sur les réseaux P2P. Notre approche exploite la technique de streaming multi-arborescence pour transporter sur chaque arbre créé une couche vidéo SVC. La Figure 5.1 illustre un exemple de notre proposition de streaming.

Dans cet exemple, le premier arbre permet la diffusion de la couche de base pour atteindre la qualité minimale SVC, à savoir : (QCIF, 7.5 fps). Le deuxième arbre permet la diffusion de la couche d'amélioration aux nœuds capables de décoder à la fois le flux de base et le flux d'amélioration pour atteindre une certaine qualité (QCIF, 15fps). Finalement, l'arbre 3 permet la diffusion de la couche d'amélioration 2 qui permet d'atteindre une qualité supérieure (QCIF, 30 fps).

Pour contrôler la qualité du flux vidéo sur ce type d'architecture, nous avons proposé les contributions majeures suivantes : (1) mécanisme de sélection des meilleurs pairs (nœuds) participant à la session de diffusion vidéo accompagné de stratégies de regroupement de pairs sous forme de réseaux overlay en fonction de différents critères (sémantique ou physique), (2) mécanisme de commutation de flux et de pairs qui permet de maintenir un niveau de QoS acceptable en cas de problème dans le réseau. Finalement, une autre contribution concerne les futurs réseaux P2P qui doivent collaborer avec le fournisseur de service (SP) pour garantir un meilleur niveau de QoS à l'utilisateur. La collaboration entre le SP et les réseaux P2P est appelé aussi P4P (provider portal for P2P applications). Dans ce contexte, nous avons proposé un nouveau concept d'allocation de ressources et d'échange de topologie entre SP et le réseau P2P pour faciliter la sélection des meilleurs nœuds.

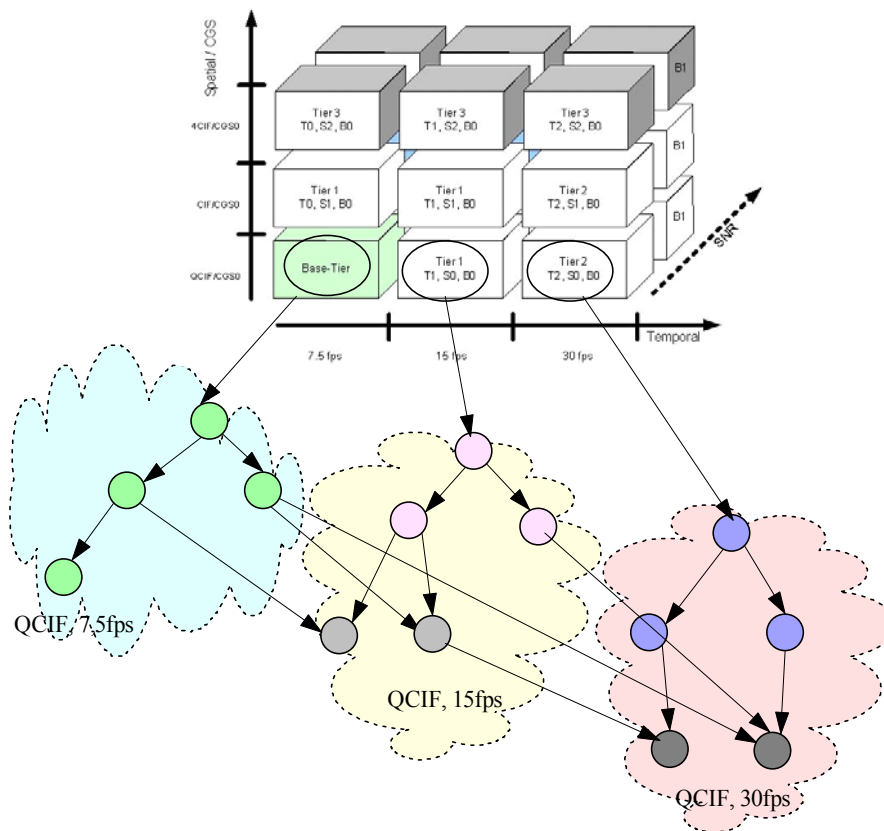


Figure 5.1 : Streaming de la vidéo SVC multi-sources sur un réseau P2P

L'architecture P2P sur laquelle nous avons développé nos contributions est de type hybride. Les pairs forment un réseau overlay structuré fondé sur l'architecture des super-nœuds (SN) et des nœuds-ordinaires (NO) pour faciliter la recherche et la sélection des nœuds. Les SNs sont généralement les nœuds qui ont le plus de bande passante, le plus de capacité disque et le plus de puissance de traitement. Toutes les requêtes leur sont envoyées. Cette structure hiérarchique nous permet de mieux supporter le passage à l'échelle pour l'implémentation de nos mécanismes. Le reste de cette section, donne plus de détails sur le fonctionnement de ces mécanismes.

5.3.1 Mécanisme de sélection des pairs

Le mécanisme de sélection des pairs a pour but de choisir parmi l'ensemble de pairs possibles ceux qui fournissent la meilleure qualité de service. On a évoqué dans la section « Etat de l'art » que les pairs peuvent être choisis sur la base de plusieurs critères physiques et/ou sémantiques. Pour cela, nous nous sommes intéressés, dans un premier temps, aux critères physiques pour la sélection des pairs, à savoir les mesures du nombre de sauts, et du RTT qui séparent le récepteur de ses émetteurs. Dans notre architecture, le nœud récepteur de flux coordonne le streaming à partir des pairs émetteurs de flux. Pour cela, il exécute un ensemble d'actions pour déterminer les meilleurs pairs disponibles en affectant un poids pour chacun d'eux (voir Eq. 5-1). Le calcul des paramètres utilisés dans l'équation 5-1 est possible grâce à des techniques de mesures actives des temps aller-retour (RTT) et du nombre de sauts (Nbre_Saut). Le poids qui en résulte est une moyenne pondérée entre le RTT et le TTL. Les simulations effectuées montrent que le fait de donner poids plus importants à RTT (exemple : $w1=0.8$, $w2=0.2$) permet de mieux estimer la qualité du lien. Même si la mesure RTT en elle-même peut suffire pour estimer la qualité de lien, il faut aussi considérer le nombre de sauts pour refléter une stratégie de choix géographique (i.e. le nœud le plus proche). Le débit TCP ou la bande passante disponible sur un lien est inversement proportionnel à la valeur RTT. Dans certains réseaux asymétriques (exemple les réseaux ADSL), la qualité du lien basée sur RTT peut ne pas convenir à la sélection des nœuds. Ceci nous a mené à considérer aussi le nombre de saut ce qui représente la proximité géographiques entre deux nœuds. Il faut noter aussi que les mesures RTT utilisées sont des moyennes exponentielles qui permettent de lisser toute variation instantanée des valeurs RTT mesurées.

$$P_i \approx \frac{RTT * w1 + Nbre_Saut * w2}{w1 + w2} \quad \text{Eq. 5-1}$$

Les meilleurs nœuds sont, par la suite choisis sur la base de leur indice de priorité (P_i). En effet, le nœud récepteur est le nœud avec un $P_i=0$ considéré comme centre de la diffusion. Les nœuds proches du récepteur selon la distance P_i sont les meilleurs nœuds (plus faible indice). Les simulations effectuées montrent qu'en utilisant ce mécanisme de priorité entre nœuds, une meilleure QoS, en terme de maximisation de la bande passante, est fournie au nœud récepteur comparé à un mécanisme ignorant les contraintes physiques du lien.

Ce mécanisme de sélection des pairs est fortement lié à la bonne perception des mesures RTT et du nombre de sauts entre les nœuds émetteurs potentiels et récepteur. Ceci peut poser un problème de passage à l'échelle (scalability). Nous avons effectué quelques tests sur Internet pour vérifier le nombre de nœuds pouvant servir un utilisateur. Les résultats sont donnés dans le Tableau 5.2.

Content	Seeds	Leechers	Total
Content 1	22370	22870	45240
Content 2	15959	10927	26886
Content 3	4261	1703	5964

Content 4	1448	738	2186
Content 5	1210	677	1887

Tableau 5.2 : Exemple de nombre de nœuds en fonction de la popularité du contenu

Il est clair que plus le contenu multimédia est populaire, plus le nombre de nœuds contenant la copie intégrale du contenu appelés « Seeds » et ceux entrain de télécharger appelés « Leechers » est grand. Pour des raisons évidentes de passage à l'échelle, notre mécanisme prend en considération uniquement un sous ensemble de « $n = 100$ » nœuds qui sont choisis sur la base du plus petit temps de réponse reçu. Parmi ces nœuds, les meilleurs nœuds sont utilisés pour la diffusion vidéo. Ainsi, pour faciliter le processus de transmission et de sélection des nœuds dans le cadre de notre approche de streaming multi-arborescence, nous avons organisé les nœuds dans des réseaux overlay. Chaque arbre de streaming représente un overlay contenant plusieurs nœuds. Cette organisation est dite hybride puisqu'elle tient compte, à la fois, de la couche vidéo SVC à transmettre (critère sémantique) et des contraintes physiques de lien en terme d'indice de priorité P_i . À l'intérieur de chaque overlay, les nœuds sont organisés suivant un arbre binaire de type « MinHeap ». Un arbre « MinHeap » binaire ou un monceau (tas) est un arbre dans lequel il existe un ordre entre un nœud et ses descendants. Dans le « MinHeap », chaque nœud a une valeur plus petite ou égale à celle de ses deux fils. Ceci nous permet d'accéder directement à la racine de l'arbre contenant le nœud avec le plus faible indice de priorité (c'est-à-dire le meilleur nœud). La Figure 5.2 illustre l'organisation en overlay hybride pour le streaming de la vidéo SVC sur un réseau P2P. L'arbre utilisé vérifie deux propriétés importantes : d'une part la valeur de chaque nœud est inférieure ou égale à celle de ses enfants, d'autre part, l'arbre est quasi complet. Dans un arbre complet, chaque ligne, exceptée peut être la dernière, contient un nombre maximal de nœuds équivalent à 2^i (i étant la profondeur dans chaque niveau, 0 pour la racine, 1 pour les enfants de la racine, etc.). Avec cette organisation, chaque overlay maximise le débit qu'il peut offrir. Rappelons que le but de cette organisation est de maximiser la capacité de réception du nœud récepteur par le choix des meilleurs pairs afin de faciliter la gestion de l'adaptation vidéo. Ainsi, la couche de base SVC, qui est considérée comme la couche la plus importante, est transmis à partir des meilleurs nœuds, et ainsi de suite pour les couches d'amélioration.

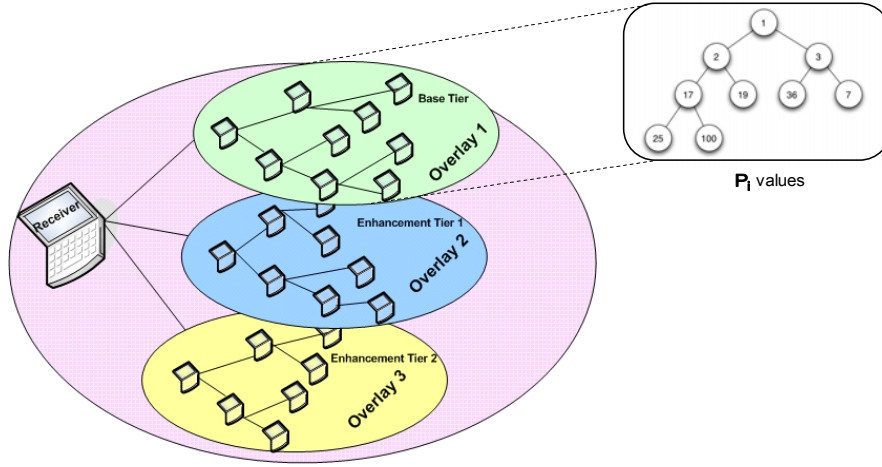


Figure 5.2 : Organisation hybride du réseau P2P pour le streaming vidéo

Chaque overlay a une capacité « upload » inversement proportionnelle à son indice de priorité qui est calculé par la somme des tous les indices des nœuds qui lui sont connectés. Le nœud récepteur choisi les « n » premiers nœuds de chaque overlay en parcourant l'arbre depuis la racine pour atteindre le débit de la couche SVC. L'estimation du débit est donnée à titre indicatif dans l'équation Eq. 5-2. Pour faire correspondre le débit du lien avec le débit de la couche SVC, le mécanisme de spécification de trafic doit être utilisé par chaque nœud émetteur. Ce mécanisme est détaillé dans la prochaine section.

$$Bandwidth = 1.3 * MTU / (RTT * \sqrt{Loss}) \quad \text{Eq. 5-2}$$

Les résultats des simulations effectuées prouvent que, grâce à ce mécanisme de sélection des meilleurs nœuds et à la gestion efficace de la mémoire tampon de réception, une meilleure QoS est fournie au flux vidéo, comparé à un streaming P2P ignorant les contraintes physiques des liens et sémantiques de la vidéo.

5.3.2 Mécanisme d'adaptation vidéo sur P2P

La variation de la qualité des liens de communication couplée au caractère dynamique des nœuds impose que le streaming vidéo soit adaptatif, dynamique, et attentif à tout changement intervenant dans le réseau P2P, aussi bien au niveau des caractéristiques physiques (RTT, nombre de sauts) que des caractéristiques sémantiques (distribution des couches SVC). Le codage vidéo SVC offre de très grandes flexibilités d'adaptation de flux vidéo pour des terminaux et des réseaux hétérogènes. SVC fournit des fonctionnalités pour déterminer la qualité d'un flux vidéo en termes de SNR, résolution temporelle et résolution spatiale. Nous avons spécifié quelques scénarios permettant à l'utilisateur de choisir les adaptations possibles pour augmenter/réduire la qualité du flux reçu. Quatre modèles d'adaptation importants ont été définis : horizontal, vertical, diagonal, et ZigZag. Dans chaque modèle, le niveau de qualité SNR peut être modifié. Chaque modèle augmente/réduit l'un des critères de

qualité vidéo (c'est-à-dire SNR, résolution spatiale, résolution temporelle) par des adaptations. La Figure 5.3 illustre les quatre modèles d'adaptations possibles qui pourront être choisis par le nœud récepteur dans sa politique d'adaptation. Il faut noter que la variation de la qualité SNR n'est pas illustrée dans cette figure.

C'est ainsi la résolution spatiale est privilégiée dans le processus d'adaptation horizontale, et ainsi de suite. Pour chaque couche SVC, une mémoire tampon est utilisée et chaque paquet, appartenant à une couche SVC, reçu par le nœud récepteur, est mis dans la mémoire tampon correspondante, avant le processus de multiplexage global.

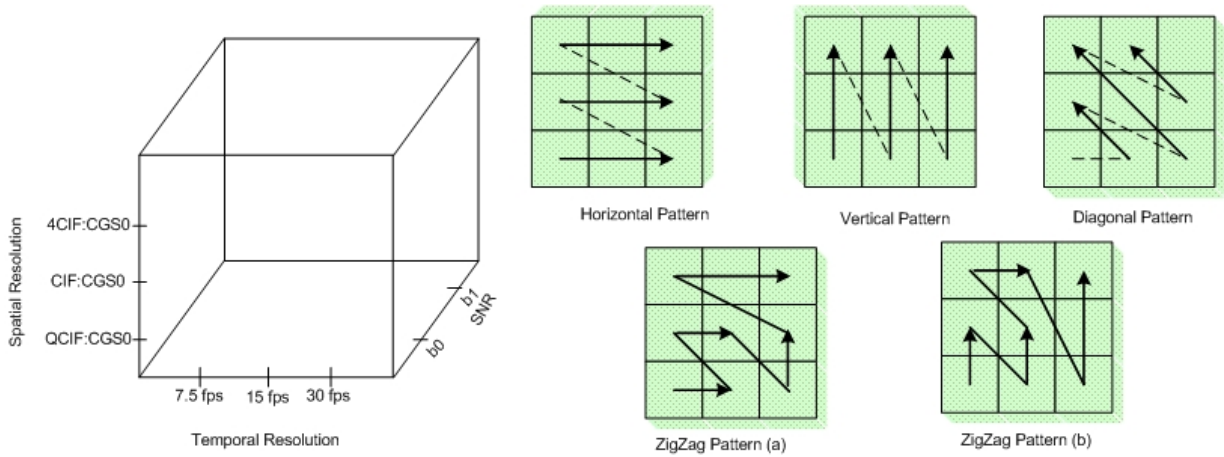


Figure 5.3 : Modèles d'adaptation des flux SVC

Chaque couche SVC doit être décrite par un modèle TSPEC (Traffic Specification) qui donne les paramètres essentiels de description du trafic à envoyer par un ou plusieurs nœuds P2P (voir le Tableau 5.3). Ces paramètres peuvent être disponibles dans le codec temps réel ou calculé sur des tranches de vidéo stockée.

Acronym	TSPEC Parameter
PDR	Peak Data Rate of video layer (kbps)
ADR	Average Data Rate of video layer (kbps)
PBS	Peak Burst Size of video packets (kb)
X_{avg}	Average Packet Size of video packets (kb)
X_{max}	Maximum Packet Size of video packets (kb)

Tableau 5.3 : Paramètres utilisés dans la description du trafic

En se basant sur le TSPEC, nous pouvons déterminer la viabilité de la réception de chaque paquet vidéo. Pour cela, deux paramètres sont utilisés : $\beta_k^{schedule}$ et $\beta_k^{receive}$, avec $\beta_k^{schedule} = \{s_k, p_k^{playback}, p_k^{transmit}\}$ et $\beta_k^{receive} = \{s_k, p_k^{playback}, p_k^{receive}\}$.

s_k représente la taille d'un paquet vidéo k , $p_k^{playback}$ représente le temps relatif de lecture du paquet k par rapport au paquet précédent, $p_k^{receive}$ est le temps de réception relatif du paquet k par rapport au paquet précédent et finalement $p_k^{transmit}$ est le temps de transmission attendu d'un paquet k par rapport au paquet précédent. La principale tâche de l'algorithme d'ordonnancement est d'assurer que les paquets vidéo arrivent bien dans la fenêtre de réception attendue pour éviter toute situation de vidage/débordement de la mémoire tampon.

Pour satisfaire les contraintes de TSPEC et avoir un débit continu à la réception, deux contraintes sont dérivées.

La première contrainte concerne le débit maximum. En effet, chaque paquet vidéo envoyé sur un intervalle de temps t_i ne doit pas casser la relation du débit maximal PDR (Peak Data Rate) et doit donc obéir à la règle

$$\text{suivante : } \forall i : \frac{s_i}{t_i} \leq PDR$$

Sur un intervalle de temps déterminé (comprenant m paquets vidéo), nous avons la relation suivante, dérivée de la règle précédente :

$$\sum_{k=1}^m s_k \leq (\sum_{k=1}^m p_k^{transmit}) PDR \quad \text{Eq. 5-3}$$

La deuxième contrainte concerne le débit moyen de réception. En effet, le processus d'arrivée généré par l'algorithme d'ordonnancement doit obéir à la condition suivante :

$$0 \leq \sum_{k=1}^m s_k - ADR \sum_{k=1}^m p_k^{transmit} \leq PBS \quad \text{Eq. 5-4}$$

Ceci permet d'éviter toute situation de vidage/débordement de la mémoire tampon.

Si $D_{\max}$ représente la limite du délai de transfert dans un réseau, alors quelque soit le groupe de m paquets vidéo la relation dans l'équation Eq.5-5 doit être vérifiée.

$$0 \leq \sum_{k=1}^m p_k^{receive} - \sum_{k=1}^m p_k^{transmit} \leq D_{\max} \quad \text{Eq. 5-5}$$

C'est la relation la plus importante pour garantir une QoS au flux vidéo. Cette relation détermine aussi la limite de la gigue pour un paquet vidéo. $D_{\max}$ peut être spécifié en se basant sur une estimation du RTT maximum toléré entre le nœud émetteur et récepteur. Dans notre cas, $D_{\max}$ est inspiré de la technique utilisée par TCP pour le calcul du temporisateur de retransmission (voir Eq. 5-6). $D_{\max}$ représente aussi la gigue maximale possible pour tous les paquets.

$$D_{\max} = \min[U_{bound}, \max[L_{bound}, (\alpha * SRTT)]] \quad \text{Eq. 5-6}$$

U_{bound} représente la taille maximale du buffer de réception en terme de délai, L_{bound} représente un délai minimum de garde qui peut être de l'ordre de 50% U_{bound} , α est le facteur de la variance du délai, il peut être entre 1.3 et 2, finalement SRTT est la valeur du RTT lissée par une moyenne exponentielle permettant de donner plus de poids aux observations du RTT récentes par rapport aux observations passées. Pour les services temps réel non interactifs tel que le streaming vidéo, la valeur de U_{bound} peut se situer autour de quelques secondes, typiquement 10 à 15s.

Chaque buffer de réception, au niveau récepteur qui correspond à un nœud émetteur, est modélisé par les paramètres suivants : bT , bA et bR . Le paramètre bT correspond à la taille du buffer en terme de temps, bA est la partie active de ce buffer, et bR est la partie de confiance (buffer de garde). Ces paramètres représentent des délais. Tant que la partie de confiance est remplie, aucune adaptation n'est nécessaire. Si la partie de confiance se vide, une adaptation doit être appliquée. La partie de confiance (délai de confiance) bR correspondante à la couche de base SVC peut être configurée de manière plus importante que les autres couches. Ceci permet d'assurer que la couche de base est toujours présente dans le buffer de réception. Si m paquets vidéo sont déjà disponibles dans la partie de confiance, alors la partie active du buffer est modélisée comme suit (Eq. 5-7) :

$$bA = \sum_{i=m+1}^n p_k^{playback} \quad \text{Eq. 5-7}$$

La condition suivante assure que le buffer ne se vide et ne déborde pas :

$$0 < \sum_{i=m+1}^n p_k^{playback} + bR \leq bT \quad \text{Eq. 5-8}$$

Si la condition de l'équation Eq. 5-7 n'est pas assurée, alors une adaptation vidéo doit être appliquée. Dans ce cas, deux politiques d'adaptation sont possibles, qui peuvent être appliquées simultanément ou l'une après l'autre. En effet, si la partie du buffer de confiance des couches d'amélioration commence à se vider, alors on a le choix de retenir un autre nœud, puisque celui-ci est devenu non fiable (nœud qui quitte le réseau, ou conditions réseaux dégradées). En choisissant un autre nœud d'émission, on doit appliquer une commutation de nœud et attacher un nouvel algorithme d'ordonnancement à ce nœud. Dans le cas où la partie de confiance se vide totalement, alors une adaptation SVC est nécessaire, c'est-à-dire qu'il faut passer d'une qualité SVC à une autre. Si la partie de buffer de confiance de la couche de base se vide, alors plus de données sont à décoder, on doit dans ce cas figure, stopper la lecture du flux et démarrer un nouveau processus de sélection de nœuds.

Les résultats des simulations effectués montrent qu'en utilisant un mécanisme d'adaptation vidéo, nous assurons une continuité de QoS pour les flux vidéo SVC.

5.3.3 Concept de collaboration P2P/Fournisseur

Une nouvelle génération de réseaux P2P est entrain de naître, celle-ci préconise une meilleure collaboration entre le réseau P2P et le fournisseur de services ou SP (Service Provider) ainsi que le fournisseur des infrastructures réseaux ou NP (Network Provider). En effet, SP et NP ont réalisé qu'ils sont entrain de perdre le contrôle sur ce qui se passe dans leurs réseaux à cause du déploiement massif des réseaux P2P. Cette situation pose aussi des problèmes aux utilisateurs P2P qui veulent télécharger des contenus rapidement et avec une certaine garantie de QoS. Nous avons présenté dans la section 5.3.1 un mécanisme permettant la sélection des meilleurs nœuds pour l'ouverture d'une session de streaming temps réel. Ce mécanisme pourrait être effectué efficacement par le SP, du moment qu'il a une maîtrise totale des ressources disponibles et de celles contractées avec d'autres fournisseurs (grâce aux SLA SP-NP et NP-NP). Le SP peut choisir, par exemple, un nœud d'émission sur la base géographique, physique ou tout simplement choisir des nœuds pour mieux contrôler les ressources utilisées. Le but est de permettre aux utilisateurs P2P de mieux consommer les contenus disponibles par une collaboration entre le réseau P2P et le fournisseur de service. Cette initiative est en cours de concrétisation dans le cadre du group de travail P4P (Provider Portal for P2P applications) [P4P]. L'IETF n'a pas échappé à ce mouvement par la création du groupe de travail ATLO (Application-Layer Traffic Optimization) [ATLO]. Le groupe de travail ATLO définit actuellement la problématique d'optimisation du trafic P2P [MAR 08] par des mécanismes qui préconise une collaboration entre le réseau P2P et les différents fournisseurs. Dans ce contexte, nous avons proposé un nouveau concept d'allocation de ressources et d'échange de topologie entre SP et le réseau P2P pour faciliter la sélection des meilleurs nœuds. La Figure 5.4 illustre le processus de déroulement d'un streaming vidéo assisté par les fournisseurs SP/NP que nous avons appelé « SP-Driven P2P network ». Nous proposons, dans cette architecture, d'utiliser le modèle de réseaux P2P hybride qui fait appel à des super-nœuds (SN). De plus, nous proposons que le SN soit désigné par le SP. Ainsi, pour chaque demande d'allocation de ressources, un nœud consommateur effectue sa requête via le SN. Ce dernier, informe le SP de la demande d'allocation du nœud récepteur. Une fois l'allocation confirmée, et la matrice de trafic est mise à jour au niveau du SP, le streaming vidéo peut commencer. Nous proposons que nœud SN fournisse trois types de services : (1) informe les autres nœuds sur les capacités du fournisseur réseau tel que les classes de services utilisées, (2) spécifie quand et comment les ressources du fournisseur réseau seront utilisées, (3) décrit la topologie du réseau tels que le numéro du AS (Autonomous System), position géographique des nœuds, etc.

avons proposé un processus d'allocation de ressources et de sélection des meilleurs nœuds à travers une communication explicite entre le réseau P2P et le fournisseur de service.

5.5 Mes Publications Associées

- [1] Mubashar Mushtaq, *Toufik Ahmed*, "IPTV Service Delivery over SP-driven P2P Networks - A Framework for Linear to Non-Linear Service Distribution" to appear in IEEE Consumer Communications & Networking Conference (**CCNC 09**), Las Vegas, 10 – 13, January 2009.
- [2] Mubashar Mushtaq, *Toufik Ahmed*, "QoS Provisioning for Video Streaming Over SP-Driven P2P Networks Using Admission Control" to appear in IEEE Consumer Communications & Networking Conference (**CCNC 09**), Las Vegas, 10 – 13, January 2009.
- [3] Mubashar Mushtaq and *Toufik Ahmed*, "P2P-based Collaborative Media Streaming for Heterogeneous Network Terminals", in proc. of IEEE International Conference on Automation, Quality and Testing, Robotics (**AQTR 2008**), pp. 330-335, May 2008.
- [4] Mubashar Mushtaq and *Toufik Ahmed*, "P2P-based Mobile IPTV: Challenges and Opportunities" in proc. of the 6th IEEE /ACS International Conference on Computer Systems and Applications (**AICCSA'08**), pp. 975-980, April, 2008.
- [5] Mubashar Mushtaq and *Toufik Ahmed*, "Multimedia Caching Integration for Efficient Content Delivery over Heterogeneous Networks" in proceedings of the 6th IEEE/ACS International Conference on Computer Systems and Applications (**AICCSA'08**), pp. 1011-1012, April, 2008.
- [6] Mubashar Mushtaq and *Toufik Ahmed*, "Smooth Video Delivery for SVC based Media Streaming over P2P Networks" in proc. of IEEE Consumer Communications & Networking Conference (**CCNC 2008**) Special session on P2P, pp. 447-451, Jan 2008.
- [7] Mubashar Mushtaq and *Toufik Ahmed*, "Hybrid Overlay Networks Management for Real-Time Multimedia Streaming over P2P Networks", in **Lecture Notes in Computer Science (LNCS)**, "Management of Multimedia Networks and Services", San José, California, USA, Vol. 4787, pp.1–13, 2007.
- [8] *Toufik Ahmed* and Mubashar Mushtaq "P2P Object-based adaptive Multimedia Streaming (POEMS)" in **Journal of Network and Systems Management**, Special Issue on Peer-to-Peer Technologies in Network and Service Management, Vol. 15, Number 3, pp. 289-310(22), september 2007
- [9] Mubashar Mushtaq and *Toufik Ahmed* "Adaptive Packet Video Streaming Over P2P Networks Using Active Measurements" in proceedings of the Eleventh IEEE Symposium on Computers and Communications (**ISCC'06**), IEEE Press, Cagliari, Italy. pp. 423 – 428, June 2006.
- [10] Djamal-Eddine Meddour, Mubashar Mushtaq, *Toufik Ahmed*, "Open Issues in P2P Multimedia Streaming", in **IEEE ICC 2006**, workshop on Multimedia Communications Workshop (**MultiCom**), Istanbul, Turkey, 11 – 15 June 2006, URL: <http://www.ieee-icc.org/2006/index/conference-program-w.html>
- [11] Mubashar Mushtaq, *Toufik Ahmed*, and Djamal-Eddine Meddour "Adaptive Packet Video Streaming Over P2P Networks" in ACM International Conference Proceeding part of International Workshop on Peer to Peer Information Management (**P2PIM**), Vol. 152, Article N°59, Hong Kong, May 2006.

5.6 Bibliographie

- [ALTO] Application-Layer Traffic Optimization (ALTO) Working Group, online <http://alto.tilab.com/>
- [ROU 04] Mema Roussopoulos, Mary Baker, et al., "2 P2P or Not 2 P2P?", in 3rd International Workshop on Peer-to-Peer systems, San Diego, CA, USA, February 2004.

- [SET] Seti@Home, Project Home Page, <http://setiathome.berkeley.edu/>
- [GEN] Genome@Home, Project Home Page, <http://genomeathome.stanford.edu/>
- [FOL] Folding@Home, project Home Page, <http://folding.stanford.edu/>
- [GRO] Groove Home page, <http://office.microsoft.com/groove>
- [DES 01] H. Deshpande, M. Bawa, and H. Garcia-Molina, "Streaming Live Media over a Peer-to-Peer Network", Technical Report Stanford 2001. Available from <http://dbpubs.stanford.edu/pub/2001-30>.
- [ZHA 07A] J. Zhang, L. Liu, and L. Ramaswamy. "PeerCast: Churn-Resilient End System Multicast on Heterogeneous Overlay Networks", Journal of Network and Computer Applications (JNCA), Elsevier, 2007.
- [BAN 02] S. Banerjee, B. Bhattacharjee, and C. Kommareddy. Scalable application layer multicast. In ACM Sigcomm, 2002.
- [CHU 02] Y.H. Chu, S. G. Rao, and H. Zhang, "A case for end system multicast" in Proceedings of ACM Sigmetrics 2002.
- [JAN 00] J. Jannotti, D. Gifford, K. Johnson, M. Kaashoek, and J. O'Toole, "Overcast: Reliable multicasting with an overlay network," Proc. the Fourth Symposium on Operating Systems Design and Implementation, 2000, pp. 197-212.
- [CAS 03] M. Castro, P. Druschel, A-M. Kermarrec, A. Nandi, A. Rowstron and A. Singh, "SplitStream: High-bandwidth multicast in a cooperative environment", SOSP'03, Lake Bolton, New York, October, 2003.
- [PAD 02] V. N. Padmanabhan, H. J. Wang, P. A. Chou, and K. Sripanidkulchai, "Distributing streaming media content using cooperative networking," in Proc. ACM NOSSDAV, Miami Beach, FL, May 2002, pp. 177-186.
- [PAD 03] V. N. Padmanabhan, H. J. Wang, P. A. Chou, "Resilient Peer-to-Peer Streaming", in IEEE ICNP, Atlanta, GA, USA, November 2003.
- [P2PCAST] P2PCast: Online: <http://www.cs.nyu.edu/~nicolosi/P2PCast>
- [P4P] P4P working group, online: http://www.dcia.info/documents/P4PWG_Mission_Statement.pdf
- [NIC 03] A. Nicolosi and S. Annapureddy, "P2PCast: A peer-to-peer multicast scheme for streaming data", Technical Report, University of New York, 2003.
- [PAI 05] V. Pai, K. Kumar, K. Tamilmani, V. Sambamurthy and A. E. Mohr, "Chainsaw: Eliminating Trees from Overlay Multicast", in proceedings of IPTPS 2005, Ithaca, NY, USA, February 24-25, 2005. Revised Selected Papers in Peer-to-Peer Systems IV, LNCS Volume 3640/2005
- [MAG] N. Magharei and R. Rejaie, "PRIME: Peer-to-Peer Receiver-driven Mesh-based Streaming", in proceedings of IEEE Infocom pp. 1415-1423, Anchorage, Alaska, May 2007.
- [ZHA 05] X. Zhang, J. Liu, B. Li, and T-S. P. Yum, "CoolStreaming/DONet: A Data-Driven Overlay Network for Efficient Live Media Streaming", In proc. of IEEE Infocom'05, Miami, FL, USA, March 2005.
- [NWM] NetWorldMap project, available at <http://www.networldmap.com>
- [VIS] Visualroute, <http://www.visualroute.com/>
- [VEN 06] V. Venkataraman and P. Francis., "Chunkyspread: Multi-tree unstructured end system multicast," in IEEE ICNP 2006, San Babara, CA, USA, 2006.

- [GAR 03] N. Garg, R. Khandekar, K. Kunal, and V. Pandit, “Bandwidth maximization in multicasting,” in Annual European Symposium on Algorithms, Budapest, Hungary, 2003.
- [CUI 04] Y. Cui, B. Li, and K. Nahrstedt, “On achieving optimized capacity utilization in application overlay networks with multiple competing sessions,” in ACM Symposium on Parallel Algorithms and Architectures 2004, Barcelona, Spain, June 2004.
- [LI 04] J. Li, P. A. Chou, and C. Zhang, “Mutualcast: An efficient mechanism for one-to-many content distribution,” in ACM SIGCOMM Asia Workshop 2004, Beijing, China, 2004
- [MAR 08] E. Marocco, V. Gurbani “Application-Layer Traffic Optimization (ALTO)”, IETF draft, Work in Progress, draft-marocco-alto-problem-statement-02 July 10, 2008
- [SMA 07] T. Small, B. Li, and B. Liang, “Outreach: Peer-to-peer topology construction towards minimized server bandwidth costs,” IEEE JSAC Special Issue on Peer-to-Peer Communications and Applications, Jan. 2007.
- [ZHA 07B] Meng Zhang, Qian Zhang, Lifeng Sun, Shiqiang Yang: Understanding the Power of Pull-Based Streaming Protocol: Can We Do Better? IEEE Journal on Selected Areas in Communications 25(9): 1678-1694 (2007)

*« Do not go where the path may lead, go instead
where there is no path and leave a trail. »
Ralph Waldo Emerson*

6 Conclusion Générale et Perspectives de Recherche

6.1 Conclusion Générale

Ce mémoire présente une synthèse de mes activités de recherche et d'enseignement, et de mes responsabilités administratives réalisées au sein du laboratoire PRISM de l'université de Versailles (1999-2004) en tant que doctorant, puis, depuis 2004, dans le laboratoire LaBRI de l'université Bordeaux 1 en tant que Maître de Conférences. Durant ces dernières années, mes activités de recherche ont porté sur l'analyse, la conception et l'évaluation des architectures de communication les plus appropriées pour la transmission et le contrôle de la qualité de service des flux vidéo transportés sur les réseaux filaires, les réseaux sans-fil et les architectures P2P. Par ailleurs, d'autres travaux de recherches ont été menés mais ne sont pas présentés dans ce mémoire. Dans la thèse de Usman Javaid nous nous étions intéressé à la gestion du routage et de la mobilité dans les réseaux PN (Personal Networks) ou réseaux personnels. Dans le projet ENTHRONE-1, nous avons défini une architecture de gestion de la QoS de bout-en-bout basée sur le concept d'un superviseur de service appelé E-IMS (ENTHRONE Integrated Management Supervisor). Au même titre, nous avons proposé une architecture de monitoring de bout-en-bout de la QoS pour les réseaux IP. Les résultats de ces travaux de recherche ont été publiés dans des conférences et revues internationales.

Mon domaine principal de recherche est celui du transport des flux vidéo sur les réseaux de nouvelle génération (NGN). Il est clair que les réseaux NGN sont de plus en plus complexes et seront bâtis sur la technologie IP comme modèle de réseau fédérateur. Le protocole IP a montré sa simplicité et sa puissance dans le déploiement de l'Internet d'aujourd'hui, malheureusement, il présente plusieurs inconvénients liés à la garantie de la qualité de service. Les modes d'accès réseaux sont considérés comme des maillons importants dans la chaîne de distribution de service. Le succès escompté des réseaux NGN sera dû, en particulier, à l'exploitation des nouvelles techniques de transmission, de codage et de modulation tel que 802.11g et son codage adaptatif ou encore le 802.11n et sa technique MIMO (Multiple In Multiple Out), à l'exploitation de nouveaux modes de routage de type ad hoc, au transport en mode maillé (Mesh) multi-sauts sans fils 802.11s et 802.16j, et aussi à l'organisation en overlay, P2P (Peer-to-Peer) et à la virtualisation des réseaux.

En parallèle à ce développement, le domaine de l'audiovisuel et du multimédia occupe désormais une place importante au sein des technologies de l'information et de la communication. La dynamique actuelle de ce secteur de recherche est plus que jamais importante. Elle peut s'expliquer par les avancées technologiques dans

le domaine de la compression audio/vidéo et la mise au point des standards tels que : MPEG-4 part 2, MPEG-4 part 10 ou H.264/AVC (Advanced Video Coding), et récemment son extension SVC (Scalable Video Coding) qui favorisent la convergence des services audiovisuels et multimédia vers la technologie IP. Il apparaît à présent que la clef de succès pour ces services réside dans la prise en compte des finalités d'usage et des besoins de l'utilisateur. Ainsi, le défi pour les services multimédia est de prendre en compte diverses caractéristiques et d'être en mesure d'optimiser les ressources utilisées et d'assurer une meilleure qualité perçue pour l'utilisateur final c'est-à-dire Quality of Experience (QoE). Par ailleurs, l'hétérogénéité présente à tous les niveaux (réseaux, terminaux, contenus et systèmes), constitue un point important qui ralentit la convergence globale des réseaux et des services. L'interopérabilité entre systèmes hétérogènes est inévitable pour le déploiement à grande échelle des tels services. D'autres verrous relatifs à cette convergence concernent la sécurité de la transmission, la prise en charge du handover (horizontal entre plusieurs domaines et vertical entre plusieurs technologies), l'auto-adaptation, etc.

Ainsi, mes contributions de recherche dans ce domaine peuvent être classées en trois grandes catégories : (1) le transport et le contrôle de la qualité de service vidéo sur les réseaux IP filaires avec différenciation de services, (2) le transport et le contrôle de la qualité de service vidéo sur les réseaux sans fil de type 802.11, et (3) le transport et le contrôle de la qualité de service vidéo sur les réseaux P2P.

Dans les différents chapitres présentés de ce mémoire, nous avons vu que la plupart des travaux de recherche existants ou menés dans ce domaine respectent le paradigme d'isolation et d'indépendance des couches protocolaires hérité du modèle de référence des réseaux ISO/OSI. La prise en charge de la sémantique des flux dans le processus de distribution de contenus vidéo est presque inexistante. Pour assurer une meilleur QoS, mon approche conceptuelle supprime cette limite en autorisant une meilleure prise en charge de la sémantique et des contraintes applicatives (QoS au niveau applicative) par les couches et les mécanismes dédiés au transport (QoS réseaux) et vis-versa. Durant les phases d'élaboration de la solution, l'objectif à atteindre est de supporter la continuité de la QoS horizontalement (à travers les acteurs de la chaîne de service) et verticalement (dans la pile protocolaire de transport) pour offrir un service de bout-en-bout supportant la QoS/QoE. Ceci nous a mené à développer et à valider des systèmes de transport de flux vidéo sur les réseaux IP filaires, sans fil et sur les architectures P2P. Ces systèmes ont été testés et validés, soit par des simulations soit par des expérimentations dont les résultats ont été publiés dans des conférences et des journaux spécialisés dans notre domaine de recherche.

6.2 Perspectives de Recherche

La convergence des réseaux et des services constitue une avancée majeure pour notre société. Ce jeune concept qui constitue les réseaux NGN, a encore de nombreux défis à résoudre. Ces NGN associeront différentes

technologies et architectures d'accès, fixes et sans fil (DVB, Wireless IP, UMTS) à travers plusieurs plates-formes matérielles et logicielles de contrôle. Le protocole IP reste la technologie de référence pour bâtir cette convergence et offrir une multitude de services (données, téléphonie, diffusion vidéo, vidéoconférence, etc.).

Mes perspectives de recherche s'inscrivent dans ce contexte. Elles traitent l'aspect d'adaptation et de contrôle de la QoS dans la chaîne de distribution des services audiovisuels afin de supporter une QoS de bout-en-bout. Le terme de QoS est assez générique puisqu'il englobe des aspects non encore traités durant les travaux passés tels que la sécurité et la mobilité des services. Le reste de cette section présente quelques unes de mes futurs axes de recherche.

6.2.1 La mobilité

L'optimisation des services multimédia destinés à des terminaux mobiles constitue un aspect important pour la fourniture de services avec QoS. La dimension de la mobilité a été partiellement intégrée à nos mécanismes de contrôle de QoS vidéo sur les réseaux sans fil. En effet, la variation du débit physique des terminaux mobiles 802.11 a été exploitée au niveau applicatif par une adaptation du débit vidéo et de la FEC. Ce mécanisme a prouvé qu'il pouvait garantir une bonne QoS grâce à la prise en compte par l'application de la connaissance du débit engendré par la mobilité. Toutefois, la gestion de la mobilité lorsque le terminal se déplace d'un réseau à un autre doit également être considérée. Dans ce contexte, le standard 802.21 [802.21] peut apporter plusieurs aspects facilitant la gestion de la mobilité. En effet, les applications multimédia peuvent exploiter les informations fournies par 802.21 MIH (Media Independent Handover) pour un meilleur contrôle de la QoS. L'interaction cross-layer, entre les événements de bas niveau générés par le MIH et l'application multimédia, permettra d'optimiser le service lors d'une phase du handover. Les fonctions de bases fournies par le MIH : les commandes, les informations et les événements rendent l'application multimédia consciente de son environnement et lui permettra d'anticiper les moments critiques lors de l'exécution du handover.

En ce qui concerne, la mobilité des sessions dans les réseaux mobiles P2P, ce concept est relativement nouveau. En effet, la session de téléchargement / streaming de contenu vidéo doit aussi rester active et suivre l'utilisateur final en dépit de sa mobilité. Cela pose plusieurs problèmes de continuité de QoS et d'adaptation des services, pour lesquels il est essentiel d'apporter des propositions et des solutions innovantes et de démontrer leur pertinence.

6.2.2 Les nouveaux modes de distribution de contenus multimédia

Les techniques de communication sans fil évoluent constamment. L'ouverture de la bande de fréquences et le faible prix d'accès à ce type de technologie a permis un développement massif des nouvelles technologies sans fil de différents types : WLANs, WPANs, WMANs, WWANs, et MANETs. Dans ce contexte, les services multimédia sont confrontés à de nombreux verrous causés, principalement, par le manque de fiabilité du canal sans fil et son partage entre plusieurs utilisateurs. C'est ainsi que les services multimédia ont besoin d'intégrer le mieux possible dans leurs architectures les progrès au niveau des protocoles MAC et de routage afin d'assurer le provisionnement de QoS. La capacité du système de transmission à s'adapter dynamiquement aux changements de conditions du canal constitue aussi un élément clé à la fourniture de QoS. Les adaptations que nous avons présentées dans nos diverses contributions (adaptation de débit, adaptation de la FEC, marquage dynamique, etc.) sont régies par des politiques (i.e. des connaissances) préalablement établies qui dictent le nouveau comportement du système à partir de son état actuel. Cependant, aucun apprentissage n'est utilisé dans nos mécanismes d'adaptation. Ces derniers peuvent être modifiés dynamiquement par un processus d'adaptation par apprentissage jusqu'à l'obtention du comportement désiré. Ce nouveau concept d'auto-adaptation ouvre le spectre à plusieurs domaines d'applications comme la gestion des réseaux, la sécurité et le transport de flux multimédia.

Par ailleurs, les réseaux de diffusion sont considérés depuis toujours comme la solution idéale aux problèmes de distribution de contenu multimédia pour les terminaux fixes et mobiles. Plusieurs réseaux de diffusion pour les mobiles : DVB-H, DMB et MediaFLO (Media Forward Link Only) sont en cours de déploiement. D'autre part, les réseaux 2.5G et 3G offrent déjà des services multimédia comme la TV mobile et la VoD, mais seulement un nombre limité d'utilisateurs pourra seulement y accéder puisque le mode unicast est utilisé. Des extensions à ces types de réseaux sont en cours pour supporter le mode diffusion (broadcast) et multicast grâce à des initiatives comme 3GPP MBMS (Multimedia Multicast/Broadcast Services) [MBMC] et 3GPP2 BCMCS (Broadcast and Multicast Services) [BCMCS]. La combinaison des réseaux de diffusion et des réseaux classiques offre, désormais, de nouveaux services et introduit de la personnalisation et de l'interactivité. Là aussi, une attention particulière est nécessaire pour intégrer plusieurs mécanismes au niveau de plusieurs technologies réseaux (uplink / downlink) afin d'assurer une meilleure QoS à l'utilisateur.

Dans ce contexte de convergence des réseaux, les opérateurs satellites veulent participer activement à ce mouvement pour intégrer des services IP dans leurs systèmes et dans leurs offres de services. L'interactivité au niveau du terminal satellite est désormais possible grâce à l'introduction de nouveaux modes de transport tel que DVB-S,S2/RCS. Cependant, comme toutes les communications sans fil, les communications satellitaires souffrent d'un certain nombre de problèmes : les interférences, les délais de transmission, et la bande passante variable et limitée. La voie retour dans DVB-S,S2/RCS se base sur un schéma de gestion de bande passante

appelé DRA (Dynamic Rate Adaptation). Lors de la variation des conditions de propagation du signal, le système peut forcer le terminal satellitaire à changer de schéma DRA (changement de codage et de modulation) pour fiabiliser les communications. Ce changement de schéma DRA engendre une modification du débit attribué à un terminal satellite. En effet, les techniques cross-layer que nous avons proposées peuvent s'appliquer à l'adaptation du service multimédia aux conditions du lien satellite. C'est dans ce cadre, que nous avons mené une étude pour le CNES (Centre National d'Etudes Spatiales) portant sur les optimisations cross-layer pour la gestion des ressources sur les réseaux satellites. D'autres pistes de réflexions pour étendre l'interaction cross-layer sur les réseaux satellites sont aussi envisagées.

6.2.3 Les réseaux de capteurs

Les réseaux de capteurs WSN (Wireless Sensor Networks) font l'objet d'une attention particulière qui s'est traduite par un nombre important de travaux de recherche sur ce sujet. L'objectif principal est de relever les défis posés par le paradigme WSN tels que : puissance de traitement limitée, capacité de communication réduite, communication multi-sauts, et hétérogénéité des capteurs déployés. Il est maintenant possible de faire des captures ubiquitaires de flux multimédia grâce à la disponibilité des capteurs d'images, de camera vidéo, de microphones, etc. Dans ce contexte est né un nouveau modèle de réseau de capteurs multimédia appelé WMSN (Wireless Multimedia Sensor Networks). Les applications WMSN sont nombreuses : réseaux multimédia de surveillance, suivi (*tracking*) d'un objectif, surveillance de l'environnement, gestion du trafic, etc. L'ensemble de ces applications nécessite la récolte et la communication des événements sous forme de flux de types images, vidéo, et audio. D'autres défis doivent être traités tels que : le contrôle d'énergie pour les communications et les traitements multimédia, la fiabilité des communications, et la gestion de la bande passante disponible pour le transport des événements et des flux multimédia, etc. Le couplage des réseaux de capteurs avec les réseaux filaires et sans fil classiques nécessite une bonne prise en charge des caractéristiques intrinsèques existantes. Les algorithmes de routage des réseaux de capteurs sans fil exigent de grandes connaissances sur tous les niveaux : niveau d'énergie, de la pile protocolaire du réseau et de l'état des capteurs voisins. Les protocoles de routages traditionnels basés sur IP ne peuvent être appliqués efficacement aux réseaux de capteurs sans fil et donc il est nécessaire de trouver de nouvelles solutions adaptées au transport des flux multimédia.

6.2.4 Les futurs réseaux P2P collaboratifs

Les réseaux P2P constituent une avancée importante dans le domaine de la société de l'information pour faciliter l'accès aisé au contenu quelque soit sa nature. Des estimations montrent que, dans les cinq prochaines années, tous les media, sans exception, pourront être acheminés sur Internet. Cette croissance est due en partie à l'utilisation des réseaux P2P. Les opérateurs de réseaux et de services commencent à réaliser l'importance de

cette technologie dans la croissance de leur secteur d'activité et dans l'augmentation du gain qui peut en résulter. Ainsi, les réseaux P2P séduisent de plus en plus les différents fournisseurs qui commencent à les intégrer dans leurs mécanismes de transport et dans leur offre de service. Cependant, il manque une réelle synergie entre les réseaux P2P et les fournisseurs. Cette synergie doit être prise en charge dans le processus de transport et dans les mécanismes d'optimisation du trafic. Ceci permettra d'offrir une meilleure qualité aux utilisateurs finaux en optimisant le transport et le contrôle des flux. Le groupe de travail IETF (ALTO Application-Layer Traffic Optimization) [ALTO] a vu le jour récemment pour la proposition d'une solution à cette problématique de collaboration réseaux P2P / fournisseurs. Parmi les éléments envisagés, le groupe ALTO recommande la définition d'un nouveau protocole de type requête / réponse pour fournir des informations essentielles pour sélectionner les meilleurs nœuds et découvrir la topologie du réseau. Dans ce contexte, un nouveau projet Européen est en cours de définition (Call 4, FP7) dans lequel j'apporte mon expertise sur la collaboration entre le réseau P2P et les fournisseurs (réseaux/services) pour la fourniture des services multimédia temps réel : IPTV et VoD [MUS 09].

6.3 Bibliographie

[802.21] IEEE 802.21 home page Online: <http://www.ieee802.org/21/>

[ALTO] Application-Layer Traffic Optimization (ALTO) Working Group, online <http://alto.tilab.com/>

[BCMCS] 3GPP2 S.R0030-A v1.0, Broadcast and Multicast Service - Stage 1, Revision A. work in progress.

[MBMC] 3GPP TS 23.246 Multimedia Broadcast/Multicast Service (MBMS); Architecture and functional description (Release 6), work in progress.

[MUS 09] Mubashar Mushtaq, Toufik Ahmed, "IPTV Service Delivery over SP-driven P2P Networks - A Framework for Linear to Non-Linear Service Distribution" to appear in IEEE Consumer Communications & Networking Conference (CCNC), Las Vegas, 10 - 13 January 2009.

Annexe : articles sélectionnés

Adaptive Packet Video Streaming Over IP Networks: A Cross-Layer Approach

Toufik Ahmed, Ahmed Mehaoua, Raouf Boutaba, *Member, IEEE*, and Youssef Iraqi

Abstract—There is an increasing demand for supporting real-time audiovisual services over next-generation wired and wireless networks. Various link/network characteristics make the deployment of such demanding services more challenging than traditional data applications like e-mail and the Web. These audiovisual applications are bandwidth adaptive but have stringent delay, jitter, and packet loss requirements. Consequently, one of the major requirements for the successful and wide deployment of such services is the efficient transmission of sensitive content (audio, video, image) over a broad range of bandwidth-constrained access networks. These media will be typically compressed according to the emerging ISO/IEC MPEG-4 standard to achieve high bandwidth efficiency and content-based interactivity. MPEG-4 provides an integrated object-oriented representation and coding of natural and synthetic audiovisual content for its manipulation and transport over a broad range of communication infrastructures. In this paper, we leverage the characteristics of MPEG-4 and Internet protocol (IP) differentiated service frameworks, to propose an innovative cross-layer content delivery architecture that is capable of receiving information from the network and adaptively tune transport parameters, bit rates, and QoS mechanisms according to the underlying network conditions. This service-aware IP transport architecture is composed of: 1) an automatic content-level audiovisual object classification model; 2) a reliable application level framing protocol with fine-grained TCP-Friendly rate control and adaptive unequal error protection; and 3) a service-level QoS matching/packet tagging algorithm for seamless IP differentiated service delivery. The obtained results demonstrate, that breaking the OSI protocol layer isolation paradigm and injecting content-level semantic and service-level requirements within the transport and traffic control protocols, lead to intelligent and efficient support of multimedia services over complex network architectures.

Index Terms—Content-based rate adaptation, Internet protocol (IP) quality-of-service (QoS), MPEG-4/7, service-aware transport protocols, unequal error protection.

I. INTRODUCTION

RAPID ADVANCES in digital video coding and networking is leading to the development of a wide range of new audiovisual services and applications. Such applications include wired/wireless videoconferencing, interactive digital TV, remote diagnosis/surgery, distance/remote sensing, process monitoring, and tele-education. Many of these applications involve the use of enabling media coding and content anal-

ysis techniques for media compression, indexing, search and retrieval.

One of the major requirements for the successful and wide deployment of such applications is the efficient transmission of multimedia content (audio, video, text, images) over a broad range of bandwidth-constrained access network infrastructures, in particular Internet protocol (IP)-based next-generation networks.¹

Audiovisual applications have adaptive bandwidth capability, but stringent delay, jitter, and packet loss constraints, which are not supported by current IP networks. Today, IP quality-of-service (QoS) mechanisms and architectures [differentiated service (DiffServ), integrated services (IntServ)] are expected to address these requirements and enable a wide spread use of real-time IP services. However, these QoS control models are not sufficient since they operate on per-IP domain, and not on an end-to-end basis. Service Level agreements (SLAs) are expected to address the end-to-end service provisioning, but in the context of mobile IP multimedia services, SLAs are hard to implement since there may not be enough resources available in some areas of the network as the terminal is moving to.

It is, therefore, important to design multimedia applications capable to adapt to system and network resource constraints while ensuring that end-user requirements are met. Our work differs from others by integrating digital information, including the content itself, semantic descriptions, QoS information, and to integrate them in one layer (cognitive layer) to enable efficient, scalable, and reliable content distribution to the end-user.

The key contribution of this paper is a combination of media content analysis techniques and network control mechanisms for adaptive video streaming over IP networks. We have designed, implemented, and evaluated a cross-layer video streaming system that includes the three following components.

- A content-based video classification model for automatic translation from video application level QoS (e.g., MPEG-4 object descriptor and/or MPEG-7 meta-data framework) to network system level QoS [e.g., IP DiffServ per-hop-behaviors (PHBs)].
- A robust and adaptive application level framing protocol with video stream multiplexing and unequal forward error protection.
- A fine grained TCP-Friendly video rate adaptation algorithm.

The proposed mechanisms work together in order to achieve seamless quality of service. Their performance evaluation is carried out using network simulator 2 (ns2).

¹Note that we do not address the entire family of next-generation networks as our scheme is targeted to wired networks mainly.

Manuscript received November 1, 2003; revised May 15, 2004.

T. Ahmed and A. Mehaoua are with the CNRS-PRISM Laboratory, University of Versailles, F-78035 Versailles, France (e-mail: tad@prism.uvsq.fr; mea@prism.uvsq.fr).

R. Boutaba and Y. Iraqi are with the School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: rboutaba@bbr.uwaterloo.ca; iraqi@bbr.uwaterloo.ca).

Digital Object Identifier 10.1109/JSAC.2004.839425

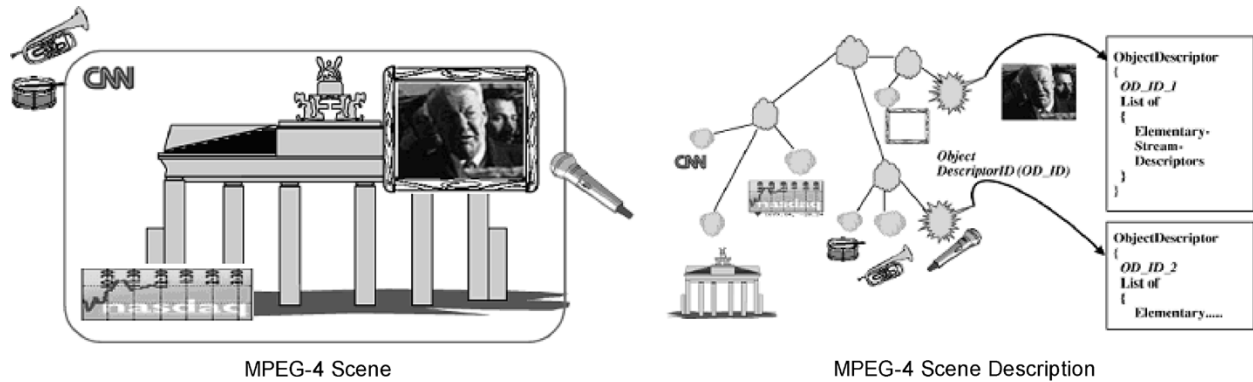


Fig. 1. MPEG-4 AVOs with corresponding object descriptors.

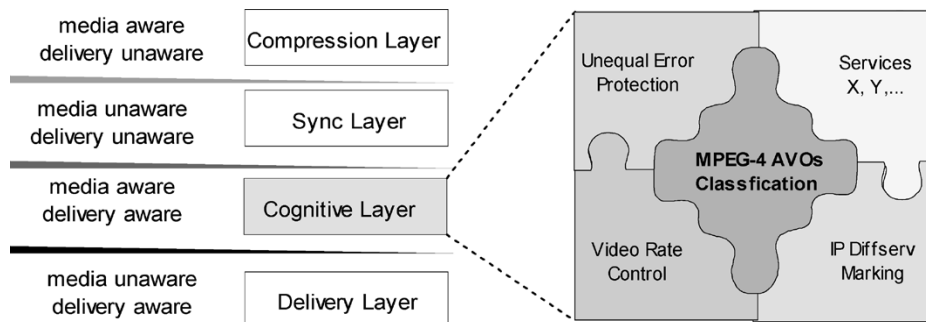


Fig. 2. Cognitive middleware for MPEG-4 content delivery and control.

The reminder of the article is as follows: Section II presents an intelligent packet video transport mechanism that is aware of application level requirement and network QoS level. This transport mechanism implements several functionalities such as an audiovisual objects classification, an unequal error protection and a TCP-Friendly video rate adaptation. Section III presents an automatic video-object classification model. Sections IV and V describe video content protection through an unequal error protection and a fine grained TCP-Friendly video rate adaptation respectively. The overall cross-layer video streaming system architecture is analyzed in Section VI. Section VII are devoted to performance evaluation and analysis. We conclude in Section VIII.

II. INTELLIGENT PACKET VIDEO TRANSPORT FRAMEWORK

Video sequences are typically compressed according to the emerging MPEG-4 multimedia framework to achieve bandwidth efficiency and content-based interactivity. The original characteristic of MPEG-4 is that it provides an integrated object-oriented representation and coding of natural and synthetic audiovisual content for their manipulation and transport over a broad range of communication infrastructures.

Fig. 1 shows a collection of MPEG-4 objects with the associated object descriptors. These audiovisual objects (AVOs) can be classified according to some attributes or metadata in the corresponding object descriptors or provided by MPEG-7 tools. These attributes can refer to the Quality of Service parameters associated with every AVO (i.e., mean/peak bandwidth, loss rate, and jitter tolerance) or related to the structural and/or

semantic description of the content like proposed by MPEG-7 or TV-anytime.

The scene description is done at two levels: the **structure level** and the **semantic level**. The structure level indicates how the scene is composed and how the AVOs are arranged in the scene in terms of both spatial and temporal locations. The semantic level is concerned with how the various streams are configured and how they have to be delivered to the user. In particular, it describes the expected QoS requirements from the network.

To implement an efficient system for object-based MPEG-4 video transmission over IP networks with QoS management capabilities, it is necessary to distinguish between important AVO and less important one. The MPEG-4 AVOs are classified based on application-level QoS criteria and AVO's semantic descriptors according to AVO's descriptors and MPEG-7 framework [1]. The classification model leads to a relative priority score (RPS) for each AVO. The MPEG-4 AVOs requiring the same QoS performance (i.e., having the same RPS) from the network are automatically grouped and associated within one of the IP DiffServ PHB. Object data-packets within the same class are then multiplexed and transmitted over the selected transport layer with the corresponding bearer capability and priority levels. Thus, we propose to extend the MPEG-4 system architecture with a novel "cognitive layer." This layer implements the automatic content-based video object classification model. In our implementation, the "cognitive layer" makes use of a neural network classification model that is transparent to the video application and the network layers.

Fig. 2 depicts the enhanced layered MPEG-4 architecture. The classical MPEG-4 architecture is composed of three layers.

The “compression layer” is media aware but delivery unaware. The “sync layer” is media unaware and delivery unaware, and finally, the “delivery layer” which is media unaware but delivery aware. By media unaware, we mean that media characteristics are not used in the functionalities provided by this particular layer. We propose to include a “cognitive layer” which is media aware and delivery aware. This layer is introduced between the “sync layer” and the “delivery layer.” The “cognitive layer” is aware of the transported media and highly aware of network QoS capabilities. The word “cognitive” refers to being aware of and making judgments about something. A “cognitive transport protocol” will be able to sense its surroundings (i.e., service requirements and underlying network conditions) and to identify the media type being delivered. Using this information, it will then be able to adapt, without user intervention, to its service’s communication constraints.

Essentially, cognitive transport protocols will be able to create their own adaptable, on-the-fly traffic control level, based on the QoS requirements of the service at that moment. This fit well in the concept proposed in [24] “to build a fundamentally different sort of network—services and applications—that can assemble itself given high level instructions, reassemble itself as requirements change, automatically discover when something goes wrong, and automatically fix a detected problem or explain why we cannot do so.” The interface between “sync layer” and “cognitive layer” is called “MPEG-4 AVO interface,” and it performs a logical MPEG-4 object identification and retrieval.

The interface between “cognitive layer” and “delivery layer” is called “MPEG-4 AVO mapping interface,” it is a logical interface at which the classified MPEG-4 AVO are mapped into various network QoS models such as IP DiffServ, IntServ, MPLS, etc., and where various functionalities can be performed such as video rate control and error protection.

As shown in Fig. 2, the cognitive middleware for MPEG-4 content delivery is build around several mechanisms and is composed of the following

- *A system-level audiovisual object classification model:* That implements an efficient transmission of object-based MPEG-4 video over IP networks with QoS management capabilities. The MPEG-4 AVOs are classified based on application-level QoS criteria and AVO’s semantic descriptors according to MPEG-4 objects descriptors and MPEG-7 meta-data. This leads to a RPS for each video packet. The MPEG-4 AVO’s (e.g., video elementary streams, video layers, video plans, audio, ...) requiring the same QoS performance (same RPS) from the network are automatically classified and multiplexed within one of the available IP DiffServ PHBs. This classification mechanism provides automatic and accurate mapping between application-level QoS metrics and content relevancy to underlying QoS-capable network services.
- *A robust and adaptive application level framing protocol with data multiplexing and unequal error protection:* The application level framing (ALF) prepares the video stream to be transmitted over the selected transport layer. In case of IP network, RTP protocol is the most suitable protocol for ALF. The Audio/Video Transport Working

Group (AVT) was formed to specify a protocol for unicast and multicast real-time transmission over IP. RTP has to be customized to the transported media. Therefore, we propose a new RTP profile for MPEG-4 media that provides video object-based unequal error protection.

- *A fine grained TCP-Friendly video rate adaptation algorithm:* Video servers perform video source rate adaptation to tackle the varying network resource conditions. Based on end-to-end feedback measurements conveyed by real-time transport control protocol (RTCP) reports, the video source can estimate the allowed transmission rate and conform to it. The originality of this contribution is to finely adjust the source video rate according to the RPS of the AVOs calculated by the classification model.

The proposed MPEG-4 cognitive layer satisfies the following features.

- *Object-based abstraction:* the cognitive layer deals with MPEG-4 AVOs as a fundamental calculation entity. This allows a flexible, extensible, scalable and simple manipulation of the MPEG-4 scene.
- *Flexibility:* The flexibility of the proposed layer is achieved by: 1) allowing different functionalities and services to plug-in it (not only the one proposed in this section) and 2) enabling or disabling on the fly some of its functionalities.
- *Extensibility:* The cognitive layer can be used to derive new elements for different domains. As an example, a content examination can be used in video surveillance system to detect and track objects within the scene.
- *Scalability:* The cognitive layer augments the scalability of the system by the classifying media object into classes according to some criteria such as visual features (e.g., size and color), semantic relevance (e.g., relevance to user interest profile, background objects, and foreground objects), service quality (e.g., media features, bit-rate, and loss ratio), and/or temporal features. For example, a terminal in the MPEG-4 scene of Fig. 1 with limited capacity can choose to view a limited video sequence without logo, background, and other banners.
- *Simplicity:* The “cognitive layer” implements a simple classification algorithm based on media QoS requirements. Additional objects and features can be easily added in a modular and flexible way.
- *Application domain:* The Cognitive layer is generic and supports a broad range of applications and media transport mechanisms.

III. VIDEO OBJECT-BASED CLASSIFICATION MODEL

A. Background

Classification has been the subject of frequent and profound investigations. It has proved a useful tool in real-world applications. In networking, packet classification can be used in network elements such as edge routers for packet forwarding and filtering [2]. In [3], the authors present algorithms that can be used for packet classification and can be categorized as basic search algorithms, geometric algorithms, heuristic algorithms,

or hardware-specific search algorithms. These algorithms are used in IP framework such as firewalls and quality of service mechanisms.

In the machine-learning area, classification methods are largely used. One well-established approach is Bayesian classification, a technique that has become increasingly popular in the recent years in part due to recent developments in learning with Bayesian belief networks [4]. Another classification method based on similarity such as K-nearest neighbors (K-NN), Naïve-Bayes and hierarchical clustering are very used in machine learning.

In the area of classification, neural networks solve many problems that conventional above methods cannot, or at least not within acceptable cost or performance constraints. In this paper, we have used neural network algorithms for automatic AVO classification.

In order to benefit from object-based compression, we propose to classify the MPEG-4 AVOs at the video server from most important AVO to least important AVO. We deal with the AVO as an independent calculation primitive. Several methods can be used for AVO's classification. During scene creation, one can affect the adequate priorities to each object in the scene. For scenes with no assigned object priorities, MPEG-4 objects descriptors and/or MPEG-7 [1] can provide the relevant information needed to compute the RPS for each objects.

The MPEG-7 standard describes generic description schemes (DSs) for image, video, multimedia, home media, and archive content. MPEG-7 aims to create a multimedia content description standard in order to facilitate various multimedia searching and filtering applications. We can use this content description to perform an intelligent classification of AVOs. The main components of the image, video, and multimedia DSs are objects, feature classification, object hierarchy, entity-relation graph, code downloading, multiabstraction levels, and modality transcoding. Each AVO may have one or more associated features, which are grouped in the following categories: media features, visual features, temporal features, and semantic features. Each feature is described by a set of descriptors.

The user can interact with the MPEG-4 server and decide at any time to choose some AVOs among several others available in the scene. This is the basic type of classification. Automatic classification is performed at the server by a prioritization mechanism which assigns a relative priority score (RPS) to each AVO. High RPS values (high priorities) are assigned to the relevant AVOs in the scene (e.g., base layer stream in hierarchical coding or foreground objects of the scene) and low RPS value are assigned to the less important AVOs (e.g., enhancement layers stream or background objects).

B. Audio Visual Object (AVO) Classification Model

We suppose that we have a collection of MPEG-4 AVOs that must be grouped into classes. The number of classes is predefined and each class has well-known characteristics. A class is viewed as a class of service of a network layer. It can be an IP DiffServ class (best effort, EF, or AFx class), an ATM class of service (ABR, VBR, UBR, etc.), or any others network service abstraction that delivers QoS.

The MPEG 4 video coding standard provides an object-based representation of the video scene by allowing separate coding of AVOs. Texture and shape coding in MPEG-4 is very similar to the coding of frames in MPEG-2. Temporal instance of a video object is called video object plane (VOP). VOP is divided into macro, luminance, and chrominance blocks. VOP supports intracoded (I-VOP) temporally predicted (P-VOP) and bidirectionally predicted (B-VOP) frames. The different components contained in the video data stream do not have the same impact on the quality of the decoded video. The damage caused by some data loss in a reference picture (I-VOP or P-VOP) will affect subsequent picture(s) due to interframe predictions. Subsequently, I-frames must be protected more than P-frames and P-frames more than B-frames. Let us consider the example of video object coded using layered wavelet transform techniques. The most important layer contains the low-frequency subband of the picture, called *base layer* (BL). Other layers, which represent a hierarchical level of resolution of the wavelet transform, are less important. These layers are called *enhancement layers* (ELs).

This is a basic classification model for preparing the MPEG-4 access unit (AU) to be transmitted over the network. It operates within a single audiovisual object. A more complete classification is handled by the cognitive layer which implements a classification model. The result of the classification is a set of AVOs sorted according to their importance in the scene. The classification model assign a final RPS to each AU to apply different mechanisms such as service differentiation, video rate control or unequal error protection. This priority score reflects both the priority of a particular AVO in the scene and the priority of a single frame type (I, P, B, or hierarchical stream if any BL or EL). In the rest of this section, we will focus on the prototype implementation using a neural network algorithm.

Let $T = \{(X, c)\}$ be a training set of N labeled vectors, where $X \in \mathbb{R}^n$ is a feature vector $X = (x_1, x_2, \dots, x_n)^t$ and $c \in T$ is its class label from an index set T (X is an AVO, and T is the set of available classes, e.g., DiffServ PHB). The variables x_i are referred to as attributes (e.g., QoS features of each AVO). A class is modeled by one or more prototype, which has $\{U_1, U_2, \dots, U_m\}$ as features.

A classifier is a mapping function called C defined as $C: \mathbb{R}^n \rightarrow T$, which assigns a class label in T to each vector in $\mathbb{R}^n$ (a vector is an MPEG-4 AVO features). Typically, the classifier is represented by a set of model parameters $\Lambda = \{\Delta_k\}$. The classifier specifies a partitioning of the feature space into regions $R_j = \{X \in \mathbb{R}^n : C(X) = j\}$, where $\bigcup_j R_j \equiv \mathbb{R}^n$ and $\bigcap_j R_j \equiv \Phi$.

It also induces a corresponding partitioning of the training set into subset T .

There are different methods based on this definition, which allow an automatic classification of vectors. The classification model used in our cognitive layer is based on radial basis function (RBF) classification.

Fig. 3 shows the RBF classifier. A vector to be classified is passed to a set of basis functions, each returning one scale value φ_j , $j = 1, 2, \dots, l$. The concrete choice of the basis function is not critical; the common implementations prefer the radial basis

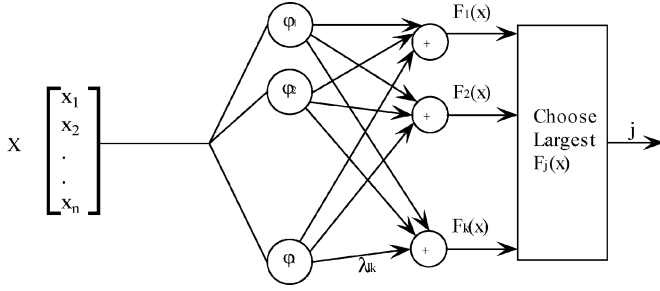


Fig. 3. RBF classification model.

function (the Gaussian “bell” functions). Hence, the name: RBF network, with: $\varphi_j(X) = e^{-(D^2/2\sigma_j^2)}$

In RBF, $D = \|X - U_k\|^2$ is the distance between X and U_k . $\varphi_j(x)$ measures similarity, i.e., the mapping function. λ_{jk} in the Fig. 3. is a set of scalar weights that connect each of the receptive fields to the class outputs of the network. The general equation of an output of the neuron j is given by: $F_k(X) = \sum_{j=1}^n \lambda_{jk} \varphi_j(X)$

The classifier maps the vector X (i.e., an AVO) to the class with the largest output: $R_j \equiv \{x \in R^n / F_j(x) \geq F_k(x) \forall k\}$

A neural network takes as input the feature vector X , produces computing class outputs $\{F_j(x)\}$, and then, classification decisions are made based on the largest output. The classification cannot be accomplished without knowing the characteristics of the underlying transport mechanisms.

IV. AVO PROTECTION THROUGH UNEQUAL ERROR PROTECTION

The proposed “cognitive layer” integrates an unequal error protection that is based on the above AVO classification model [25]. Error resilience of each *Elementary Stream* associated to one AVO can be enhanced when sensitive data are protected, whereas the subjectively less relevant audiovisual information are none or less protected during transmission, as investigated in, [5]–[9]. The IETF work in progress described in [10] and [11] specify how error protection is unequally applied to different parts of the video stream. We adopt this approach with adaptation to object based video coding streams (i.e., MPEG-4). Consequently, the classification model specifies how to assign priority score to each AU within an AVO.

A. Reed–Solomon (RS) Codes

Reed–Solomon (RS) codes aims to produce at the sender n blocks of encoded data from k blocks of source data in such a way that any subset of k encoded blocks suffices at the receiver to reconstruct the source data [12]. RS code is called an (n, k) code. RS code (n, k) is defined over the Galois Field $\text{GF}(2^q)$, where each block contains q bits. The codeword length n is restricted by $n \leq 2^q - 1$. We choose q to be 8 bits and, therefore, $n \leq 255$. With this value for q , encoding and decoding are processed easier.

Let $x = x_0 \dots x_{k-1}$ be the source data, G an $(n \times k)$ generator matrix of the (n, k) RS code, and y the encoded data. Then, y is given by

$$y = G \cdot x. \quad (1)$$

G consists of two parts. The first part is the $(k \times k)$ identity matrix I_k . The second part is an $(h \times h)$ matrix, with $h = n - k$. G is given by (2).

When G is used as generator matrix, the blocks of encoded data include a verbatim copy of the source. It simplifies the reconstruction of source data when few losses are expected

$$G = \begin{bmatrix} 1 & 0 & \dots & 1 \\ 0 & 1 & \ddots & \vdots \\ \vdots & & & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \dots & 0 & 1 \\ (n-1)^1 & (n-2)^1 & \dots & (n-h)^1 \\ (n-1)^2 & (n-2)^2 & \dots & (n-h)^2 \\ \vdots & \dots & \dots & \vdots \\ (n-1)^h & (n-2)^h & \dots & (n-h)^h \end{bmatrix}. \quad (2)$$

B. Object-Based Unequal Error Protection Using Adaptive RS Codes

We proposed an object-based unequal error protection (UEP) for handling MPEG-4 audiovisual stream according to its importance. Let us consider U_i , the i^{th} Access Unit in the AVO of priority score p [25]. The main challenge of the proposed object-based UEP is to determine the values n_i and k_i of RS code in such a way that the (n_i, k_i) RS code is efficient. The value k_i is defined as the number of packets in which an access unit U_i is divided when no error protection is performed. The value n_i depends on the priority score of a particular AVO. It depends also on the length m_i of U_i so that the traffic overhead introduced by redundant data does not become excessive.

Once the efficient (n_i, k_i) RS code is found the transmission process starts. Data of U_i is placed in k_i horizontal packets $(S_1, S_2 \dots S_k)$. Each of them has the same packet size. Padding is added to the last packet if necessary. Then, the (n_i, k_i) RS code is applied across these packets, vertically as introduced in [13]. So, redundant data for each AVO is calculated by $h_i = n_i - k_i$. The RS code for whole AVO's objects is represented as $(n_1, k_1), (n_2, k_2), \dots, (n_L, k_L)$, where $(n_1 - k_1) > (n_2 - k_2) > \dots > (n_L - k_L)$. This means that $h_1 > h_2 > \dots > h_L$, where L is the number of AVO's object present in the scene or to be transmitted in the network when the TCP-Friendly adaptation is performed.

It is clear that the proposed object-based UEP increases the amount of the traffic sent in the network. Traffic overhead is defined by the accumulated number of data-bytes in each AVO. Each AVO j with a priority score p generates traffic overhead R_{jp} given by (3) during the streaming interval (the streaming interval for the particular object is defined in the next

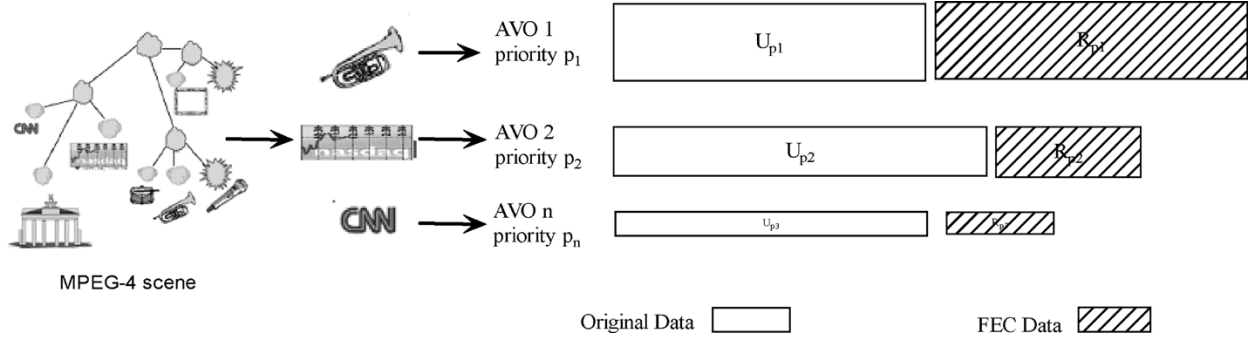


Fig. 4. RS codes for object-based unequal error protection.

section in order to comply with the TCP-Friendly adaptation requirements)

$$R_{jp} = \sum_{i=k_p+1}^{n_p} \frac{U_{ijp}}{k_{ijp}}. \quad (3)$$

Then, the accumulated overhead of the MPEG-4 scene can be denoted as

$$R(\text{scene}) = \sum_{j=1}^L \sum_{i=k_p+1}^{n_p} \frac{U_{ijp}}{k_{ijp}}. \quad (4)$$

Fig. 4 shows how RS code could be applied to MPEG-4 scene composed of different AVOs with different priority.

We assume that for each AVO j with a priority score p , error protection byte-budget B_{jp} is reserved for error protection. This byte-budget is fixed with a ratio r among the useful data U_{jp} . Then, B_{jp} can be defined as

$$B_{jp} = U_{jp} * r \quad (5)$$

and the accumulated error protection byte-budget for the MPEG-4 scene is given by (6)

$$B(\text{scene}) = \sum_{j=1}^L U_{jp} * r. \quad (6)$$

For example, with an MPEG-4 AVO of priority score $p = 1$, we can choose the ratio $r = 0.1$ then the error protection byte-budget will be limited to 10% of the useful data of this object.

In order to find the efficient value of n_i and k_i , we must satisfy two conditions. First, the traffic overhead R_{jp} is upper-bounded by the error protection byte-budget B_{jp} . The second is that k_i cannot be greater than n_i for a particular AVO. Combining (4) with the above two conditions result in a constrained optimization problem given as follows:

$$\begin{aligned} \text{Minimize : } & \sum_{j=1}^L \left| U_{jp} * r - \sum_{i=k_p+1}^{n_p} \frac{U_{ijp}}{k_{ijp}} \right| \\ & \text{for each AVO } j \text{ of a priority score } p \text{ and with } k_i \leq n_i. \end{aligned} \quad (7)$$

With the proposed object-based unequal error protection mechanism, the RS code evolves dynamically so that the network bandwidth is correctly controlled according to the network conditions.

V. FINE-GRAINED TCP-FRIENDLY VIDEO RATE ADAPTATION ALGORITHM

The majority of multimedia applications perform over an RTP/UDP/IP stack. However, UDP offers no congestion control mechanism and, therefore, is unaware of network condition and unfair toward other competing traffic. Today's Internet traffic is dominated by transmission control protocol (TCP). TCP uses several mechanisms to handle network congestion such as: additive increase and multiplicative decrease (AIMD), slow start, congestion avoidance, fast retransmit, and fast recovery. Thus, it is crucial that user datagram protocol (UDP) traffic performs also TCP-Friendly congestion control [15].

Congestion control helps preventing the application entering congestion collapse in which traffic demand is high but little useful throughput is available. Such situation can be prevented if all applications implement a TCP-like congestion control mechanism. Traffic that does not perform in a TCP-Friendly manner can be dropped by the router [16].

In our congestion control [26], the server performs rate adaptation through the adjustment of the number of streamed objects based on network state and RPS of each objects. We use a TCP-Friendly approach to adapt the server rate to network condition. The server attempts to deliver the maximum number of AVO that can fit in the current available bandwidth slot. This flow control mechanism is combined with a DiffServ marker. The server must be aware of each audiovisual object in the scene so as to classify these objects in a hierarchical manner, from less important object to more important one. As presented in Section III, it allows the server to: 1) deal with network congestion by stopping the streaming of less important objects when congestion is detected and 2) prioritize the transport of important objects through a content-aware IP packet marking scheme in IP DiffServ networks. When network congestion occurs less important AVOs will be dropped automatically by network elements. Lost packets notify the server to reduce its transmission rate by stopping streaming less important AVOs.

The idea of TCP-Friendly transport protocol is to emulate TCP behavior without replicating the TCP mechanism itself. By definition, a flow is said to be TCP-Friendly or TCP-compatible, if its arrival rate does not exceed the arrival rate of a conformant TCP implementation in the same circumstances [17]. Many TCP-Friendly congestion control mechanisms were developed recently that are either window-based or equation-based including: rate adaptation protocol (RAP) [19], loss-delay

based adaptation algorithm (LDP) [18], a spectrum of TCP-Friendly algorithm [20] and TCP-Friendly rate control protocol (TFRC) [21]. While these protocols and others are comparable in their features of simulating TCP behavior, TFRC seems to be more robust and was recently adopted by the IETF. TFRC provides sufficient responsiveness by taking into consideration all the parameters that affect the TCP rate such as loss, round-trip time (RTT) and retransmission timeout value. The key advantage of TFRC is that it has a more stable rate during the session lifetime.

Our video quality adaptation mechanism is based on TFRC. It operates as follows.

- The receiver measures the loss rate and feeds this information back to the sender. This is achieved by a modified version of RTP and RTCP protocols [14]. Each RTP packet has a timestamp and a sequence number that allow the receiver to compute the packet loss rate and the sender to compute the RTT.
- The loss rate and the RTT are then fed into the TFRC module to determine the appropriate transmission rate [see (8)].
- The sender then adds or drops audiovisual objects and the associated layers if any, to adjust its transmission rate to match the target rate (i.e., allowed rate).

The target transmission rate is computed by using the following TFRC equation [21]:

$$R_{TCP} \cong \frac{s}{RTT \sqrt{\frac{2bp}{3}} + t_{RTO} \left(3 \sqrt{\frac{3bp}{8}} \right) p(1 + 32p^2)}. \quad (8)$$

Where R_{TCP} is the target transmission rate or the allowed transmission rate, s is the packet size, RTT is the round-trip time, p is the loss rate, t_{RTO} is the TCP retransmission timeout value and b is the number of packets acknowledged by a single TCP acknowledgment.

A. Video Object-Based Rate Adaptation Algorithm

Let S be a set of MPEG-4 AVOs containing n AVO's O_j , with $j \in \{1, 2 \dots n\}$. Without loss of generality, we assume that these objects are sorted in a decreasing order of priority score carried-out through classification. Each object O_j may consist of m_j layers ($m_j = 1$). Note that lower layers within an object have higher priorities than higher layers.

Let P be the function that returns the RPS of a particular object or layer. Without loss of generality, we assume that

$$\begin{aligned} \forall j, 1 \leq j < n : P(O_{j+1}) &\leq P(O_j) \\ \forall j, 1 \leq j < n, \forall l, 1 \leq l < m_j : P(L_{j,l+1}) &< P(L_{j,l}) \\ L_{j,l} &\text{ is the Layer number } l \text{ of the Object } O_j. \end{aligned} \quad (9)$$

Using (9), we can construct an audioVisual Entity set called E composed of all object layers ordered according to their respective priorities

$$E = \{L_{1,1}, L_{1,2} \dots L_{1,m_1}, L_{2,1}, L_{2,2} \dots L_{2,m_2}, \dots, L_{n,1}, L_{n,2} \dots L_{n,m_n}\}.$$

We will note E as follows:

$$E = \{e_1, e_2, \dots, e_w\} \text{ with } w = |E| = \sum_{j=1}^n m_j.$$

Our object-based video rate adaptation mechanism operates as follows. The server evaluates the network state from the information gathered (i.e., RTT and loss rate) at time t_i , then computes the allowed sending rate R_{TCP} using (8). The server tries to send as many audio visual entities as possible without exceeding R_{TCP} taking into consideration entity's priorities. Details of the adding and the dropping processes will be presented in Sections V.B and V.C respectively.

B. Adding Audiovisual Objects

The server adds a new audiovisual entity as soon as the target rate exceeds the current sending rate of current entities plus the new entity. Assume that the server is streaming k entities at time t_i . Assume also that the client has sufficient resources to play all the entities being sent by the server. At time t_{i+1} the server can add a new entity while the following condition remains satisfied

$$\sum_{j=1}^{k+1} R_{i+1}(e_j) \leq R_{TCP}. \quad (10)$$

At the client side, the new audiovisual entity must be buffered and synchronized to the current playback time.

C. Dropping Audiovisual Objects

When the estimated throughput of the TCP session indicates that the video server is transmitting more data than it should, the later must reduce its sending rate by dropping one or more audiovisual entities. The server drops entities as long as the following condition is satisfied:

$$\sum_{j=1}^k R_{i+1}(e_j) > R_{TCP}. \quad (11)$$

D. GOV-Driven Stability

Since the TFRC compute the new target rate each RTT, adding and dropping audiovisual entities can lead to undesired bit rate oscillation and consequently video quality degradation at the receiver. To prevent from such behavior, several measures are taken.

First, the TFRC module reduces bit rate oscillations through exponentially weighted moving average (EWMA) to detect out-of-control states. EWMA statistics are used to respond dynamically to the changing values in the measured RTT and loss and regulate these values to reflect as much as possible the real network conditions. In TFRC, the loss rate is measured in terms of loss interval which represents the number between two consecutive loss events [20]. The mechanism reacts too strongly to single loss events and ensures that allowed sending rate do not change aggressively.

Second, we propose to synchronize server transmission rate changes with beginning of the group of VOP (GOV). Thus, the

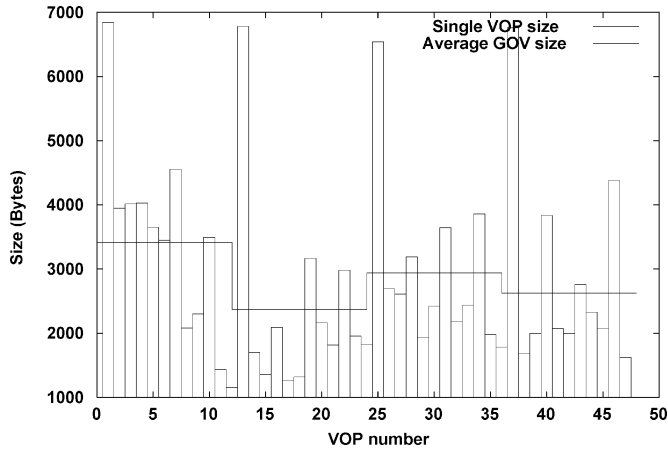


Fig. 5. Handling stability through VOP control.

new transmission rate obtained from TFRC module is used to intelligently adapt video sending rate to conform to media object encoding dependencies. Fig. 5 shows four GOVs (each group has twelve VOPs). The average line in the Figure shows the server transmitting rate at the beginning of each GOV of a current video object (VO). If this value does not fit in the current available bandwidth then the server does not stream the object.

E. Rate Adaptation and Admission Control (AC) Issues

Admission control (AC) is used to control the network load by restricting access to the network and, hence improving the level of QoS guarantee. Admission control approaches can be categorized in a number of ways such as parameter-based approaches versus measurement-based approaches and edge/end-point admission control versus hop-by-hop admission control.

Parameter-based approaches (as in [33]) assume some traffic pattern and try to maintain the aggregated resource consumption below the total capacity. They often lead to conservative resource allocation and low network utilization especially in the case of bursty traffic. In turn, measurement-based admission control relies on the measurement of current network load and, therefore, responds faster to the network status and consequently improves the network utilization [28]–[33].

As being done in the IntServ/RSVP architecture, admission control is traditionally performed on a hop-by-hop basis. Each intermediate network element along the path has to decide whether the new request can be accommodated or not and reserves resources accordingly. However, adding admission control functionality to all the core elements violates the DiffServ principle of leaving the core simple. End-point/edge admission control that pushes the admission control functionality to the edge of the network seems more suitable in a DiffServ environment and has a number of advantages over the hop-by-hop approach such as faster response time and less implementation overhead. Studies also show that simple admission control algorithms based on estimated or measured network status are generally robust [27], [28]. Most of the approaches in this category use probing [29], [32], [33] or explicit congestion notification (ECN) [27], [29] to convey the network status back to the end points. Admission decision is then made based on this feedback information. For example, in [30], state

information is embedded inside the packet header (dynamic packet state) so that admission control decisions can be made based on this information. Our proposed scheme can be used in conjunction with most of the admission control approaches proposed for DiffServ.

VI. ADAPTIVE CROSS-LAYER VIDEO STREAMING SYSTEM

Fig. 6 depicts the general block diagram of the proposed adaptive cross layer video streaming system. It is composed of a video server and a video client. The server streams the audiovisual object to the client via an IP DiffServ network using the RTP protocol. The client decodes and composes the original MPEG-4 scene. As shown in Fig. 1, each AVO is coded separately so the decoding process decodes also each AVO separately and then the composition module composes the original scene. The target transmission rate of the video server is calculated by the end-to-end feedback algorithm which is based on RTP and RTCP. This feedback information is sent to the “add/drop module,” which adapts the video transmission rate to conform to the available bandwidth. The IP DiffServ marker module handles the marking of the different RTP packet with DiffServ code point before entering the DiffServ network. The higher is the packet importance, the lower is the DiffServ drop precedence assigned by marker to this packet.

DiffServ object prioritization aims to favorite the transport of some AVOs compared with others. When network congestion occurs, less important AVO’s streams are dropped automatically by the active queue implemented in the DiffServ router. Recall that an MPEG-4 scene contains several MPEG-4 AVOs grouped according to their importance in the presentation. The IP DiffServ marker tags each video data packet belonging to one AVO with one of the supported DiffServ class of service to reflect object priority. Hence, important objects will be marked with a low drop precedence to guarantee a minimum loss, and so on.

VII. PERFORMANCE EVALUATION

A. System and Network Models

Simulations are conducted using the network simulator *ns2*. We used the network architecture shown in Fig. 7 to simulate a unicast service provided by the MPEG-4 server attached to the node “S.” The server sends data to the client attached to the node “C” using TFRC module to adapt the number of transmitted AVOs. The client is also an *ns2* agent which extends the capabilities of the RTP sink by reporting statistics information to the server. The network is loaded by n FTP streams carried over TCP (n ranges from 0 to 8). This allows the link between the routers “R1” and “R2” to be congested differently. FTP sources always have a packet to send and always send a maximum-sized (1000-bytes) packet as soon as the congestion control window allows them to do so. FTP sink immediately sends an ACK packet when it receives a data packet. The queue in the routers has a size of 50 packets. We configure two scenarios: 1) Scenario A: four FTP sources and (2) Scenario B: eight FTP sources. FTP sources send data from time $t = 30$ s until $t = 90$ s. The core IP DiffServ router examines incoming packets and reacts according to the marking, whereas “R1” is an edge router that implements marking/classification policy

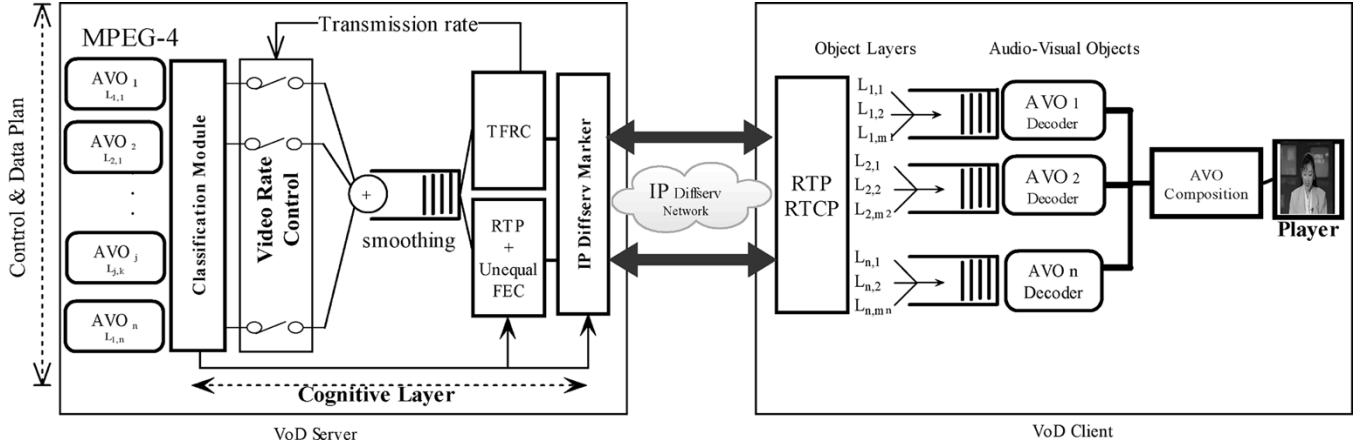


Fig. 6. General block diagram of the proposed cross-layer video streaming system.

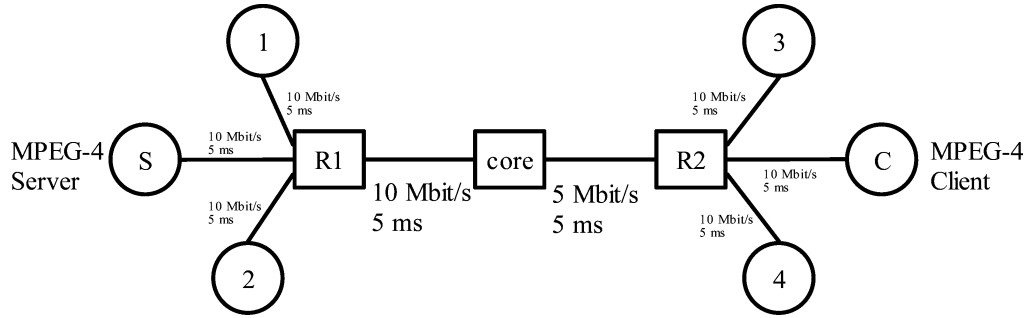


Fig. 7. Network topology for congestion control.

on incoming packets. R1 uses a two-rate three-color marker (TR3CM) [22] to mark the background. Therefore, background traffic is evenly distributed among the different DiffServ classes. Recall that the video traffic is marked at the MPEG-4 server according to AVO's priorities. The bottleneck link between the core router and R2 has a capacity of 5 Mb/s.

Some of the metrics used to setup the DiffServ domain are taken from [23] and are listed in Table I. These metrics are: 1) Priority: our configuration of the priority score of the DiffServ classes (higher value corresponds to higher priority); 2) IPTD: IP packet transfer delay; 3) IPDV: IP packet delay variation; 4) IPLR: IP packet loss ratio; 5) IPER: IP packet error ratio; and 6) BW: Bandwidth used. These values are computed statistically among the simulated DiffServ domain.

In our simulation, the MPEG-4 presentation was obtained using a set of AVO's components. We simulate the "Akiyo" video sequence shown in Fig. 8 by using four multimedia objects: AO (audio speech), VO1 (background), VO1 (speaker), and VO3 (logo). These objects are sorted as follows.

Fig. 9 shows the bit-rate of the MPEG-4 video objects that can be sent from the MPEG-4 server to the client during a period of 120 s. The complete scene is shown in Fig. 9(a). The audio object is an adaptive audio coding (AAC) MPEG-4 audio with a constant bit rate throughput of 64 Kb/s. An audio packet is sent each 125 ms. Video object 1 (the background object) has an average throughput of 200 Kb/s and a peak rate of 956 Kb/s. This object is composed of three Layers: base layer (BL), enhancement layer 1 (EL1), and enhancement layer 2 (EL2). The throughputs of the different layers are shown in

TABLE I
IP DIFFSERV QoS CLASS DEFINITIONS

Network Performance Parameter	QoS Classes				
	EF	AF11	AF12	AF13	Best Effort
Priority	4	3	2	1	0
IPTD	100ms	150ms	200ms	300ms	Unspecified
IPDV	50ms	50ms	Unspecified	Unspecified	Unspecified
IPLR	$1 \cdot 10^{-3}$	$1 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$3 \cdot 10^{-3}$	Unspecified
IPER	$1 \cdot 10^{-4}$				Unspecified
BW	1.5Mbits	1.5Mbits	1.5Mbits	1.5Mbits	4Mbits

Fig. 9(b). Video object 2 (the speaker object) has an average throughput of 650 Kb/s and a peak rate of 1722 Kb/s. This object is composed also of three layers: BL, EL1, and EL2. The throughputs of the different layers are shown in Fig. 9(c). Video object 3 (the logo object) has an average throughput of 124 Kb/s and a peak rate of 356 Kb/s. It is composed of one single layer [see Fig. 9(c)]. Table II summarizes the application level requirement for each MPEG-4 AVO.

B. Results Analysis

Intensive simulations are performed, each time with different parameters, to see the behavior of our cross-layer video streaming system.

First, we run our classification algorithm with the parameters described in the previous section, and we compute the output of each neuron. The system chooses the class of service with the largest neuron output which reflects the high similarity. We also choose the value of $\lambda_{1k} = \lambda_{2k} = \lambda_{3k} = 1$

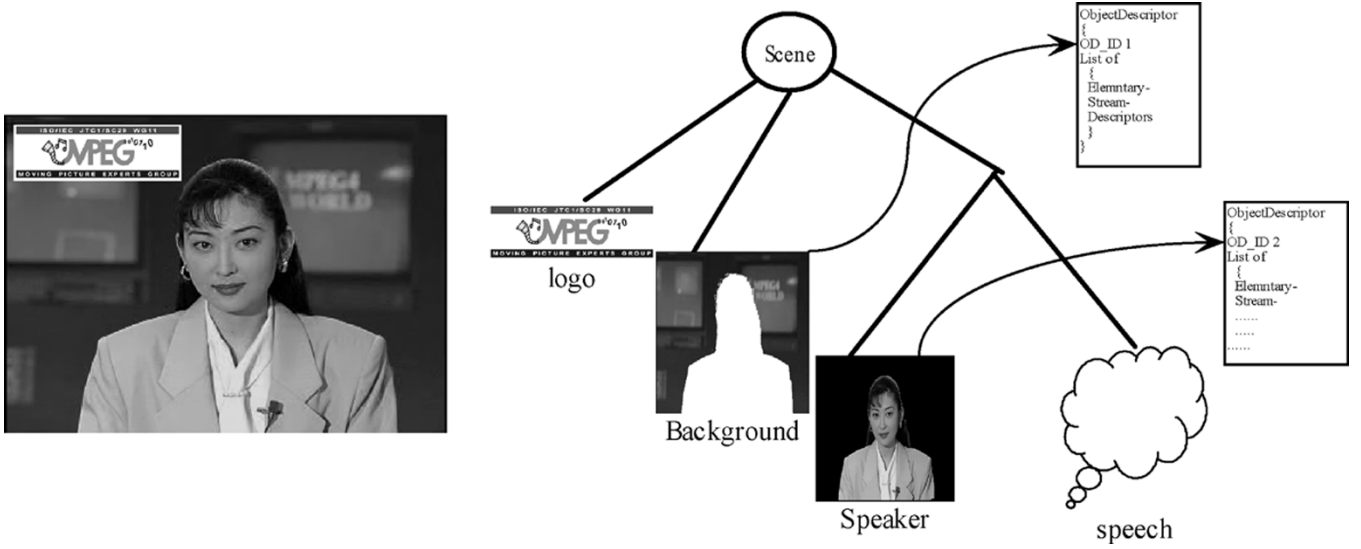


Fig. 8. Simple composite MPEG-4 scene based on “Akiyo” video sequence.

which connects each feature from the vector to the network. The feature vector of each MPEG-4 AVO is described as follows: $X = (MAX_DELAY, MAX_JITTER, LOSS_PROB)^T$ with $X \in \mathbb{R}^3$, and we have taken the value of

$$MAX_JITTER = MAX_DELAY - PREF_MAX_DELAY.$$

Output result of the RBF network (i.e., $F_k(X) = \sum_{j=1}^n \lambda_{jk} \varphi_j(X)$) is summarized in Table III. As previously explained, we select the largest output from the network. According to these results, we can see that O1 is marked with AF11 PHB and that O2, O3 are both marked with AF12 PHB. We have also measured the execution time of the algorithm and found to be about 30 ms of the mapping of 100 AVOs is about 30. The complexity of this algorithm is $O(n)^2$.

According to these results, the audio will be marked with EF PHB and has the priority score of 4. The background, and the speaker will be marked with AF11 PHB and have the priority score of 3. Finally, the logo will be marked with AF13 PHB and has the priority score of 2.

The rest of the section presents some QoS measurement namely, the video server throughput as a function of network state, packet loss, and user perceived quality.

1) *Video Server Throughput*: The video server regulates its transmission rate to comply with the allowed rate by adding or dropping audiovisual entities (video objects and video layers). Results obtained for different scenarios are shown in the figures below. Also, to simplify the interpretation of the results, Table IV summarizes the transmission ratio per AVO stream observed during the period of the simulations (120 s). Note that the FTP sources begin data transmission at time $t = 30$ s, and stop at time $t = 90$ s. VO3 has the lowest ratio since it has the lowest priority score in the scene. VO1 and VO2 have the same priority score, so the corresponding layers have more or less the same transmission ratio.

Scenario A (i.e., MPEG-4 scene with four FTP streams) is interesting since we see the effect of our adaptation mechanism.

In Fig. 10, we can see that the audio object is always present and that less important objects (respectively object layers) are not transmitted when the shared bandwidth is not sufficient. Our adaptation mechanism begins transmitting data from most important audiovisual entity to least important one. We can see that all the streams (FTP and video) fairly share the bandwidth. The logo object is not transmitted every time, because it is the less important object. When there is a lack of bandwidth in the link, the server stop streaming lower priority objects.

Scenario B (i.e., MPEG-4 scene with eight FTP streams) confirms the previous result and shows the effect of our adaptation mechanism as presented in Fig. 11. A minimum of QoS is guaranteed by our adaptation mechanism. The network does not enter in congestion since the video server is aware of network condition. Some snapshots of the real video are presented later.

2) *Packet Loss and Unequal FEC Performance*: Fig. 12 shows packet loss for scenarios A and B, using our congestion control scheme. In scenario A, we observe some loss of lower priority packets but in scenario B high priority packet are also dropped. This is due to: 1) lower priority packets are not transmitted because our congestion control mechanism regulates the server transmission rate by stopping the streaming of lower priority packets and 2) AVO1 and AVO2 require more bandwidth in our scene and cause congestion.

FTP packet loss observed in the same scenarios is also presented in Fig. 12. FTP packets encounter more loss than the video packets due to two factors. First factor is that FTP traffic is marked using TR3CM marker which distributes the marked traffic among the different classes of the DiffServ network. It is worth noting that the majority of dropped packets are those marked with high drop precedence. Second factor is that FTP source does not regulate the traffic by computing the allowed transmission rate rather it uses window-based congestion control mechanism. In case of video traffic, the damage caused by some data loss in some reference picture such as I-VOP or P-VOP will affect subsequent picture(s) due to inter-frame predictions. For example, when the I-VOP is lost, the whole

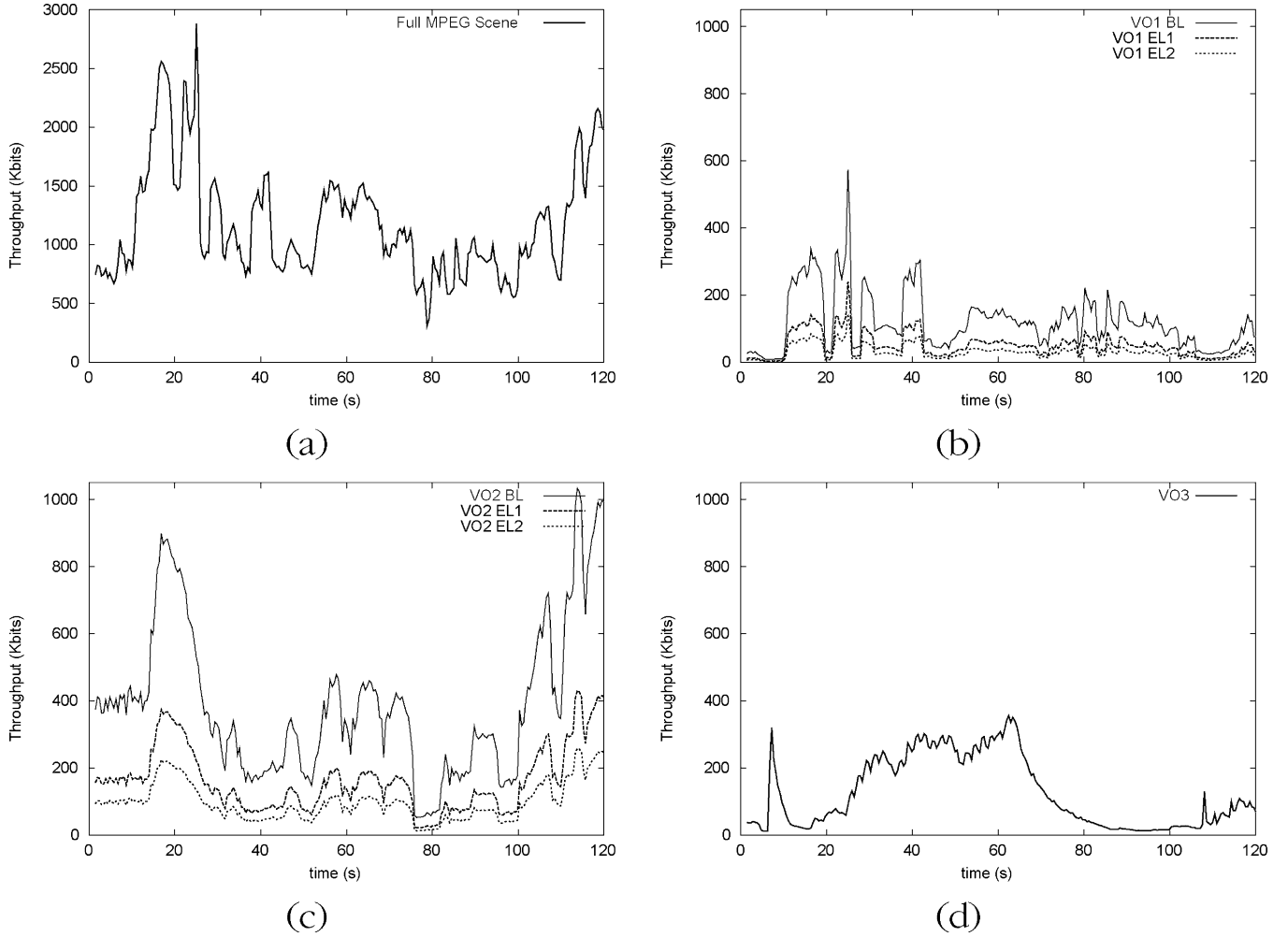


Fig. 9. Instantaneous throughput of the different MPEG-4 video objects.

TABLE II
APPLICATION-LEVEL QoS REQUIREMENTS SPECIFIED
IN MPEG-4 QoS DESCRIPTORS

QoS Metric	Synthetic Description	Audio Object	Video Object-1	Video Object-2	Video Object-3
MAX_DELAY	Absolute maximum end to end delay for the stream	150ms	250ms	250ms	350ms
PREF_MAX_DELAY	Preferred maximum end to end delay for the stream	100ms	200ms	200ms	300ms
LOSS_PROB	Allowable loss probability of any single AU.	$1 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$2 \cdot 10^{-3}$	$3 \cdot 10^{-3}$
MAX_GAP_LOSS	Maximum allowable number of consecutively lost AUs	Unspecified	Unspecified	Unspecified	Unspecified
MAX_AU_SIZE	Maximum size of an AU	342 bytes	1024bytes	1024bytes	1024bytes
AVG_AU_SIZE	Average size of an AU	342 bytes	1024bytes	1024bytes	1024bytes
MAX_AU_RATE	Maximum arrival rate of AUs	Unspecified	Unspecified	Unspecified	Unspecified
PRIORITY	Priority for the stream	Unspecified	Unspecified	Unspecified	Unspecified

TABLE III
OUTPUTS FROM THE RBF CLASSIFICATION MODEL

	EF	AF11	AF12	AF13	Best
audio	3.0	2.9962	2.9984	2.9943	2.8143
background	2.9962	2.9999	2.9995	2.9998	2.8248
speaker	2.9962	2.9999	2.9995	2.9998	2.8248
Logo	2.9927	2.9994	2.9978	2.9999	2.8273

TABLE IV
TRANSMISSION BIT-RATE RATIO PER MPEG-4 OBJECTS

Scenario	AVO	Audio	VO1			VO2			VO3
			BL	EL1	EL2	BL	EL1	EL2	
A	100%	100%	94%	87%	100%	96%	92%	55%	
B	100%	89%	60%	53%	97%	77%	71%	26%	

dependant P-VOP and B-VOP cannot be decoded. The same conclusion is valid for hierarchical streams. Hence, enhancement layer 1 cannot be decoded without the reception of base layer, and so on. When using our congestion control mechanism, lower priority audiovisual entities (those marked with high drop precedence) are not transmitted by the server when the allowed

rate decreases. This helps to prevent a future drop by the router. So, the regulation is done at the server and demonstrates clearly the advantage of our mechanism.

In order to highlight the efficiency of the unequal error protection (UEP) scheme, we compute the number of AUs

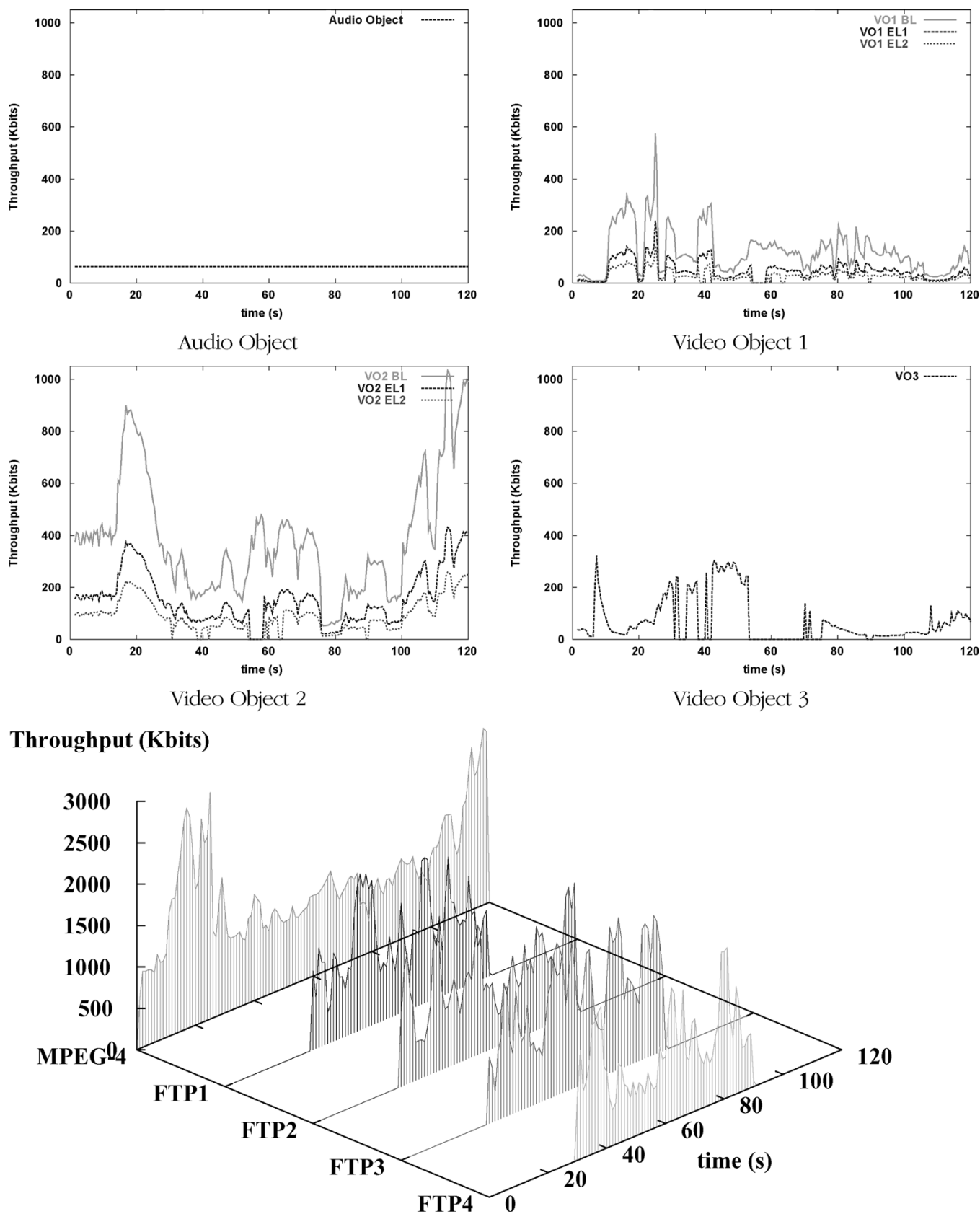


Fig. 10. Traffic throughput and fairness in scenario A.

that can be decoded at the client. Fig. 13 shows the comparison between the decoded object ratios. The X axis represents the throughput of the background traffic, and Y axis repre-

sent decoded object ratio. As expected, the quantity of AVOs decoded at the receiver side decreases when the network load increases because it entails more packet losses.

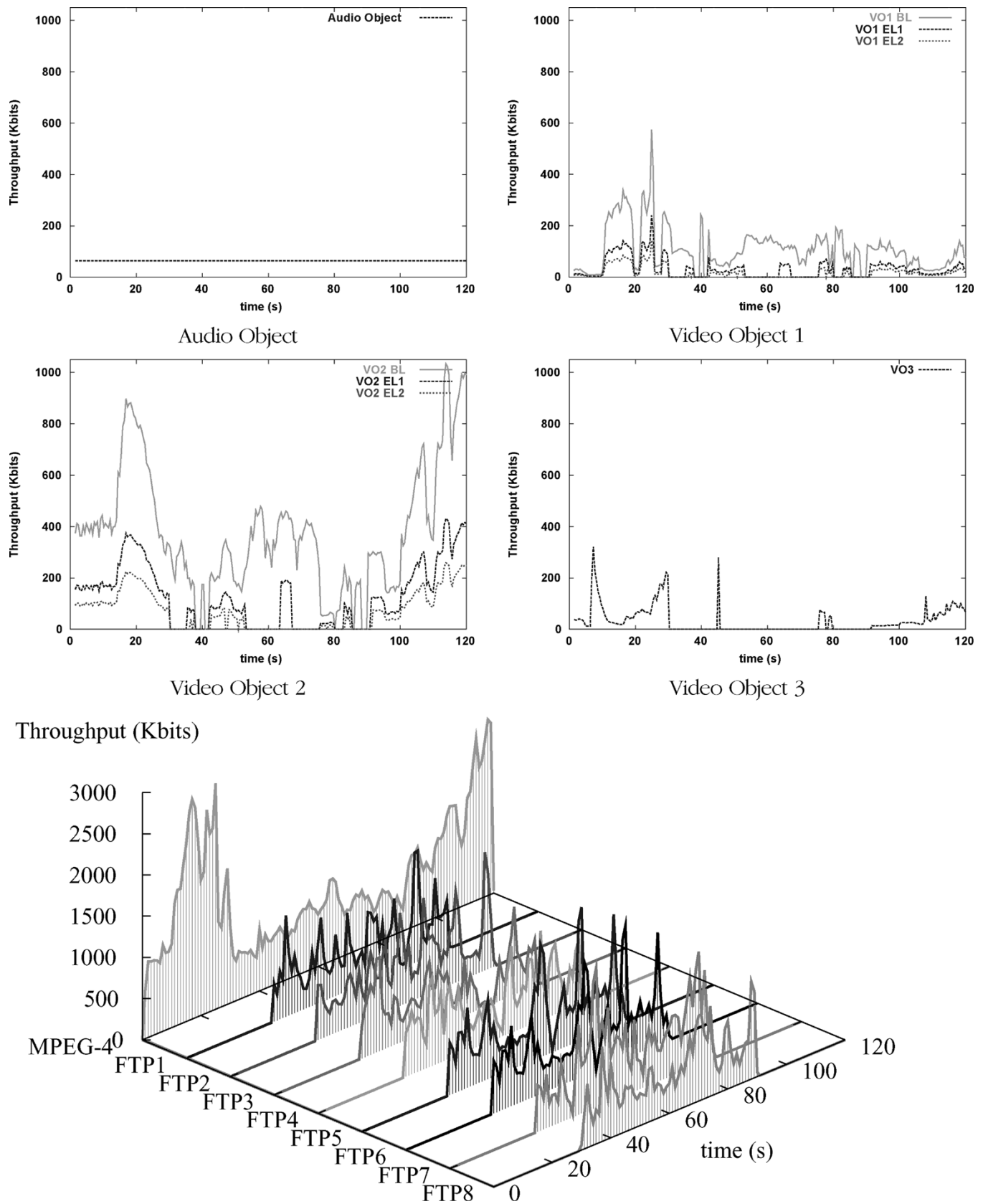
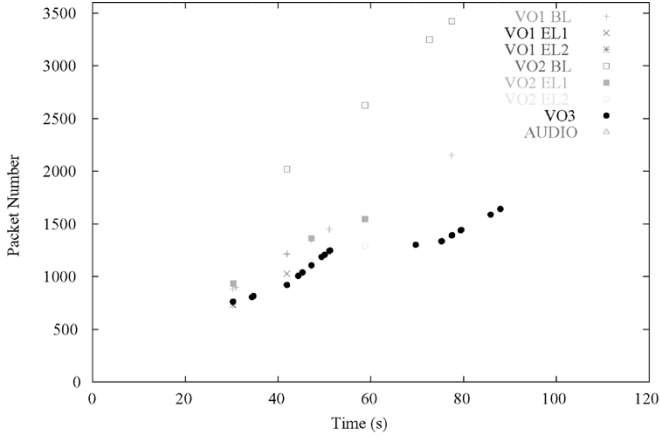


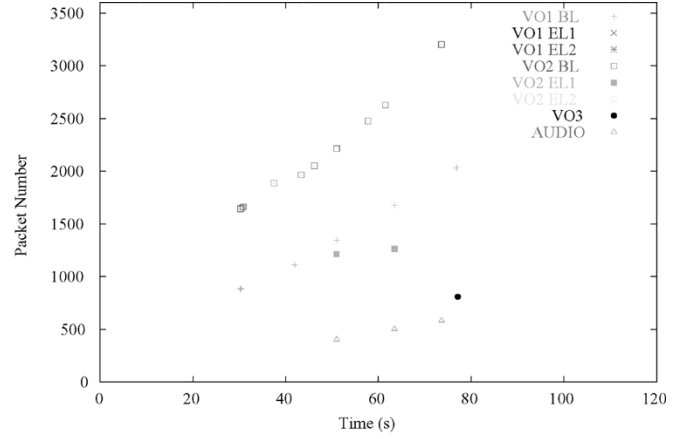
Fig. 11. Traffic throughput and fairness in scenario B.

By the use of the FEC-based UEP, packet loss rises due to an increase of the MPEG-4 packet-stream throughput by 7%. Consequently, we notice a slight higher packet losses with UEP

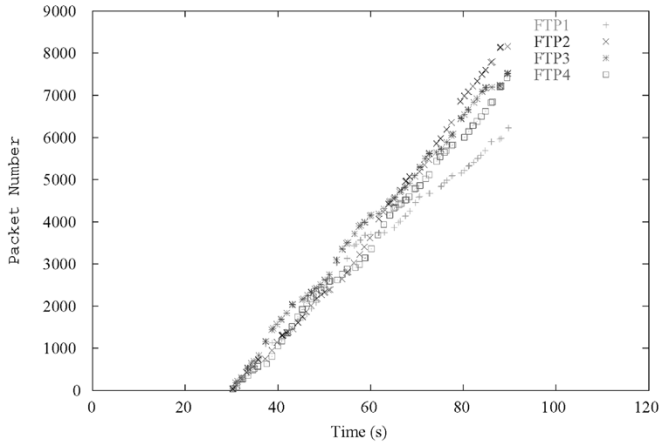
configurations, for a given network load. However, the redundant UEP information better recovers lost packet at the receiver. Consequently, some *access unit* can be restored correctly.



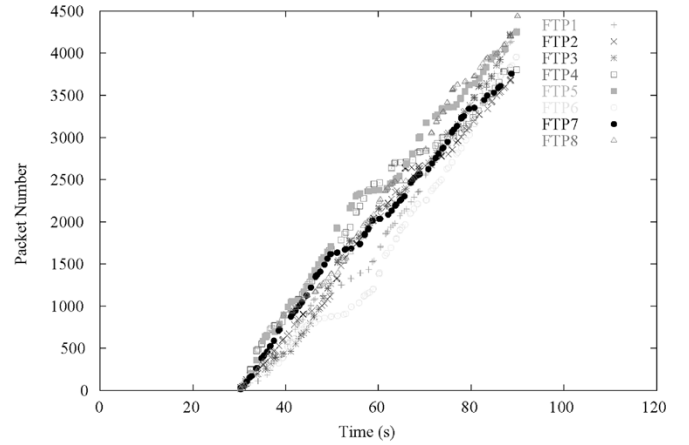
Scenario A – MPEG-4 video



Scenario B- MPEG-4 video



Scenario A- FTP traffic



Scenario B- FTP traffic

Fig. 12. Instantaneous packet loss.

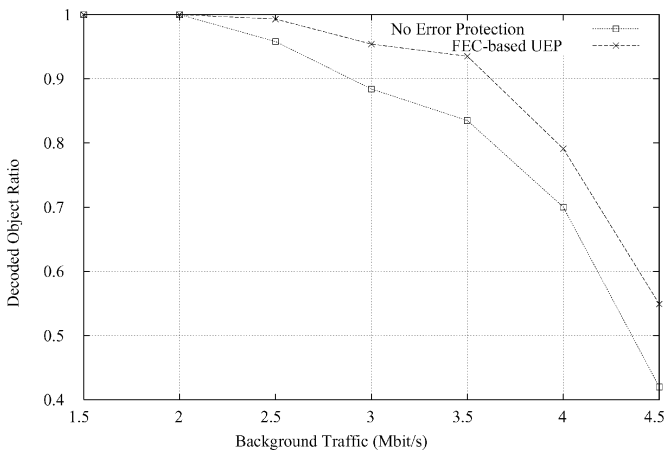


Fig. 13. Correctly decoded AV object ratio versus background traffic throughput.

Failures in the decoding process are rather distributed toward the less important objects, and then UEP reduces the effects of spatial and temporal errors propagation.

3) *User Perceived Quality*: Peak signal-to-noise ratio (PSNR) is a commonly used for measuring picture quality

degradation. It is derived from the root mean squared error. The PSNR for a degraded $N_1 \times N_2$ 8-bit image f' from the original image f is computed according to the formula in (12)

$$\text{PSNR} = 20 \log_{10} \frac{255}{\sqrt{\frac{1}{N_1 N_2} \sum_{x=0}^{N_1-1} \sum_{y=0}^{N_2-1} [f(x, y) - f'(x, y)]^2}}. \quad (12)$$

In order to compute the PSNR value, we should rebuild the received scene. By using ns2 trace file, we create the received MPEG-4 scene. The measured PSNR indicates the difference between the original and the received video sequences. Fig. 14 shows comparison between the original and the received scene quality for scenarios (A and B). In scenarios A and B, the degradation of the received quality is due to our adaptive delivery system which only sends pertinent audiovisual objects to the client. This PSNR measurement does not reflect the semantic of what the client received. Because in our case, the logo object is basically not sent. This affects the PSNR value but has no importance for the end user (in our assumption). Fig. 15 presents snapshots of the received video in scenario B.

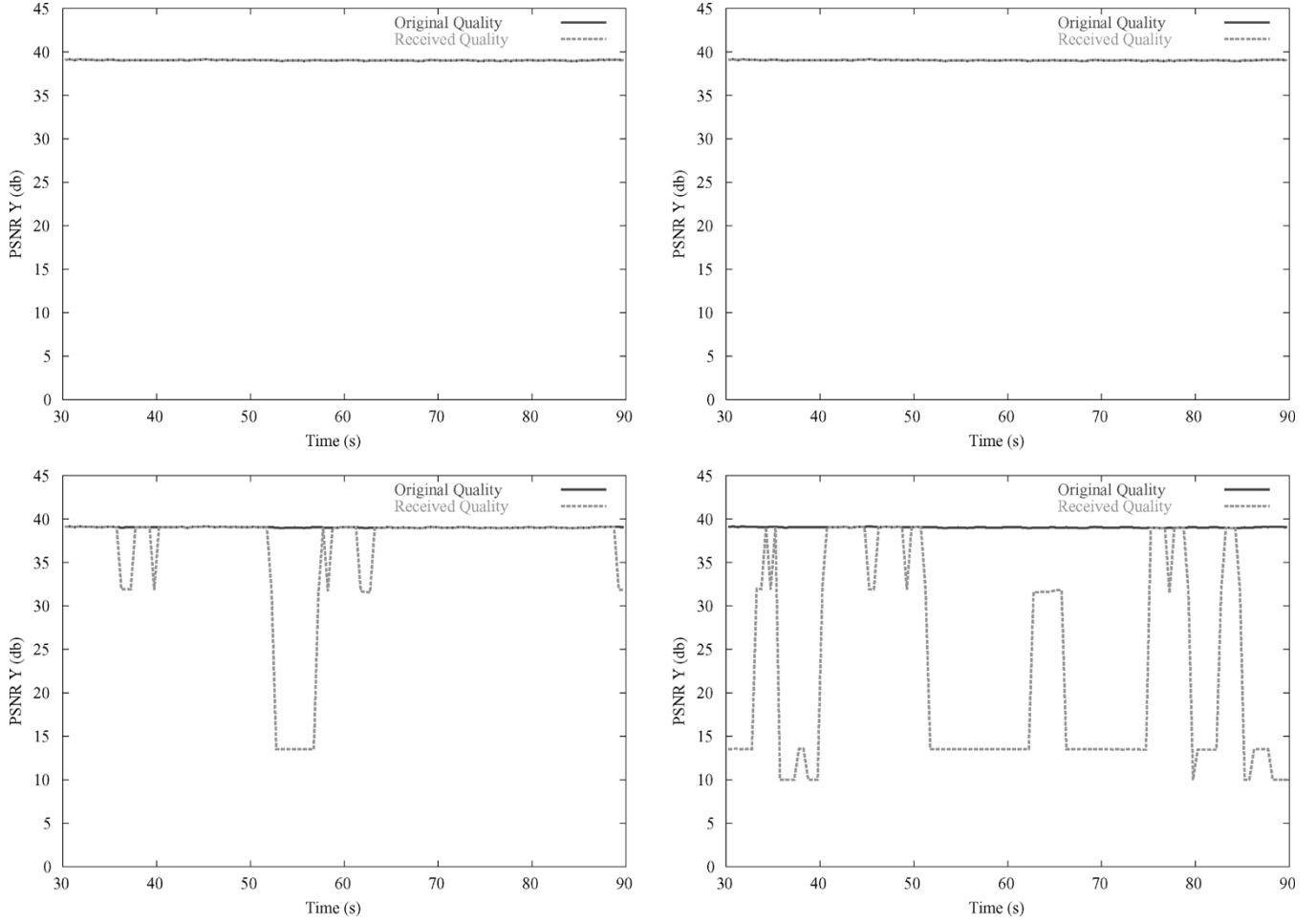


Fig. 14. Perceived video quality (PSNR) measurements for the different scenario.

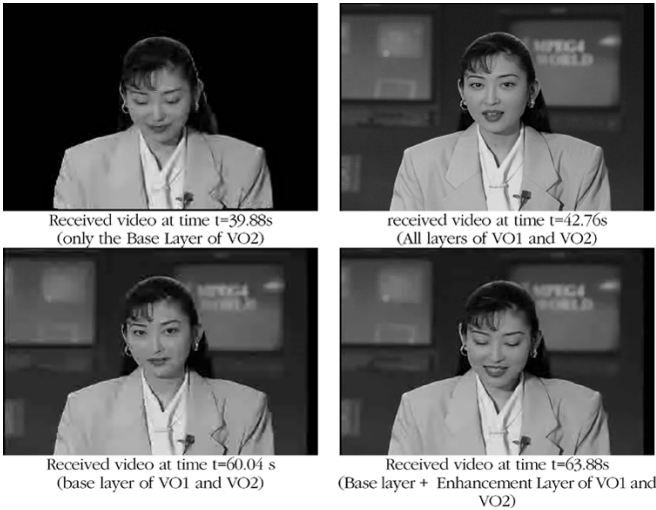


Fig. 15. Received video presentation.

VIII. CONCLUSION

In this paper, we proposed a cross-layer video streaming system that integrates an innovative “cognitive layer.” This “cognitive layer” implements intelligent services and is capable of interfacing the underlying network technology to provide: 1) a system-level automatic audiovisual object classification

model; 2) a robust and adaptive ALF protocol with fine-grained TCP-Friendly rate control and unequal error protection; and 3) an application-level video packet marking algorithm to be deployed on DiffServ-enabled networks.

This “cognitive layer” is an extension to the MPEG-4 system architecture that makes the use of a neural network classification model to dynamically and accurately group audiovisual objects of a scene with same QoS requirements to create elementary video streams that are subsequently mapped to IP DiffServ PHBs. These MPEG-4 AVOs are classified based on application-level QoS descriptors and MPEG-7 content-descriptive metadata. Thus, MPEG-4 AVOs requiring same QoS from the network are automatically classified and multiplexed within one of the IP DiffServ PHB. Object data-packets within the same class are then transmitted over the selected transport layer with the corresponding bearer capability and RPS.

The transmitted MPEG-4 streams take also benefit from the proposed “cognitive layer” by applying an unequal error protection according to the priority score of each object. The amount of recovered data is related to audiovisual objects priority score in the MPEG-4 scene. The more the object is important, the more the recovered data is valuable and better.

For fair share of bandwidth and higher user perceived quality, we have designed a content-based rate adaptation mechanism

for MPEG-4 video streams that uses a TCP-Friendly Rate Control. The novelty of our mechanism is to perform video rate adaptation by adding and dropping MPEG-4 AVOs according to their subjective relevancy to the service (i.e., through MPEG-4/7 object descriptors/metadata analysis) and instant network congestion estimations.

We have evaluated the performance of the proposed IP video streaming system by simulation using ns2. The MPEG-4 server implemented in ns2 uses the TFRC module as an equation-based congestion control mechanism. We coupled end-to-end congestion control with a DiffServ network that guarantees objects prioritization within the network. The simulation results show that subjectively important multimedia streams (audio, video in the foreground) are preserved from loss by the routers in situation of network congestion. These results show clearly the gain obtained in term of visual quality compared with a classical video streaming system without a "cognitive layer."

Combining these control mechanisms into a coherent and cognitive middleware demonstrates the usefulness and gains of interfacing content-level semantic and service-level QoS requirements to the underlying next-generation network technologies.

REFERENCES

- [1] "MPEG-7 Overview," ISO/IEC JTC1/SC29/WG11, N4980, 2002.
- [2] S. Iyer, R. R. Kompella, and A. Shelat, "ClassiPI: An architecture for fast and flexible packet classification," *IEEE Netw. (Special Issue on Fast IP Packet Forwarding and Classification for Next-Generation Internet Services)*, vol. 15, no. 2, pp. 33–41, Mar./Apr. 2001.
- [3] P. Gupta and N. McKeown, "Algorithms for packet classification," *IEEE Netw.*, vol. 15, no. 2, pp. 24–32, Mar./Apr. 2001.
- [4] D. Heckerman *et al.*, "Real-world applications of Bayesian networks," in *Commun. ACM*, vol. 38, Mar. 1995, pp. 24–26.
- [5] J. Rosenberg and H. Schulzrinne, "RTP payload format for generic forward error correction," Request for Comments, IETF, RFC 2733, 1999.
- [6] J.-C. Bolot and T. Turletti, "Adaptive error control for packet video in the internet," in *Proc. IEEE ICIP*, Lausanne, Sep. 1996, pp. 232–239.
- [7] G. Carle and E. W. Biersack, "Survey of error recovery techniques for IP based audiovisual multicast applications," *IEEE Netw.*, vol. 11, no. 6, pp. 24–36, Jun. 1997.
- [8] I. Rhee and S. Joshi, "Error recovery for interactive video transmission over the internet," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 1033–49, Jun. 2000.
- [9] B. Girod, K. Stuhlmüller, M. Link, and U. Horn, "Packet loss resilient internet video streaming," in *Proc. SPIE Visual Commun. Image Processing*, San Jose, CA, Jan. 1999, pp. 833–844.
- [10] G. Liebl, M. Wagner, J. Pandel, and W. Weng. (2003) An RTP payload format for erasure-resilient transmission of progressive multimedia streams. [Online]. Available: draft-ietf-avt-uxp-06.txt
- [11] A. H. Li *et al.* (2003) An RTP payload format for generic FEC with uneven level protection. Internet Draft. [Online]. Available: draft-ietf-avt-ulp-07.txt
- [12] S. Lin and D. J. Costello, *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall, 1983.
- [13] A. Albanese, J. Bloemer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Trans. Inf. Theory*, vol. 42, no. 11, pp. 1737–1747, Nov. 1996.
- [14] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A transport protocol for real-time applications," Request for Comments, IETF, RFC3550, 2003.
- [15] M. Mahdavi and S. Floyd, "TCP-Friendly unicast rate-based flow control," Tech. Note sent to the end2end-interest mailing list, 1997.
- [16] S. Floyd, M. Handley, J. Padhye, and J. Widmer, "Equation-based congestion control for unicast applications," in *Proc. ACM SIGCOMM*, 2000, pp. 43–56.
- [17] S. Floyd and K. Fall, "Promoting the use of end-to-end congestion control in the internet," *IEEE/ACM Trans. Netw.*, vol. 7, no. 4, pp. 458–472, Aug. 1999.
- [18] D. Sisalem and H. Schulzrinne, "The loss-delay adjustment algorithm: A TCP-Friendly adaptation scheme, network and operating system support for digital audio and video," in *Proc. NOSSDAV*, Cambridge, U.K., Jul. 8–10, 1998, pp. 215–226.
- [19] R. Rejaie, M. Handley, and D. Estrin, "RAP an end-to-end congestion control mechanism for real-time streams in the Internet," in *Proc. IEEE INFOCOM*, vol. 3, New York, Mar. 1999, pp. 1337–1345.
- [20] S. Jin, L. Guo, I. Matta, and A. Bestavros, "A spectrum of TCP-Friendly window-based congestion control algorithms," *IEEE/ACM Trans. Netw.*, vol. 11, no. 3, pp. 341–355, Jun. 2003.
- [21] M. Handley, S. Floyd, J. Padhye, and J. Widmer, "TCP-Friendly rate control (TFRC): Protocol specification," Request for Comments, IETF, RFC 3448, 2003.
- [22] J. Heinanen and R. Guerin, "A two rate three color marker (TRTCM)," RFC 2698, 1999.
- [23] "Internet protocol communication service—IP performance and availability objectives and allocations," ITU-T, ITU-T Recommendation Y.1541, 2001.
- [24] D. D. Clark, C. Partridge, J. C. Ramming, and J. T. Wroclawski, "A knowledge plane for the Internet," in *Proc. ACM SIGCOMM*, Aug. 2003, pp. 3–10.
- [25] T. Ahmed, A. Mehaoua, and V. Lecuire, "Streaming MPEG-4 audiovisual objects using TCP-Friendly rate control and unequal error protection," in *Proc. IEEE Int. Conf. Multimedia Expo*, vol. 2, Baltimore, Jul. 2003, pp. 317–320.
- [26] T. Ahmed, A. Mehaoua, R. Boutaba, and Y. Iraqi, "IP video streaming with fine-grained TCP-friendly rate adaptation," in *Lecture Notes in Computer Science*. New York: Springer-Verlag, 2003, vol. 2839, IFIP/IEEE Management of Multimedia Networks and Services, (MMNS'03), pp. 18–31.
- [27] A. Nafaa, T. Ahmed, and A. Mehaoua, "Unequal and interleaved FEC for wireless MPEG-4 video multicast," in *Proc. IEEE Int. Conf. Commun.*, vol. 27, Paris, France, Jun. 2004, pp. 1431–1435.
- [28] F. P. Kelly, P. B. Key, and S. Zachary, "Distributed admission control," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 12, pp. 2617–2628, Dec. 2000.
- [29] L. Breslau, E. Knightly, S. Shenker, I. Stoica, and H. Zhang, "Endpoint admission control: Architectural issues and performance," in *Proc. SIGCOMM 2000*, Stockholm, Sweden, 2000, pp. 57–69.
- [30] I. Stoica and H. Zhang, "Providing guaranteed services without per flow management," in *Proc. ACM SIGCOMM*, Cambridge, MA, Sep. 1999, pp. 81–94.
- [31] R. Mortier, I. Pratt, C. Clark, and S. Crosby, "Implicit admission control," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 12, pp. 2629–2639, Dec. 2000.
- [32] B. Pang, H. Shao, W. Zhu, and W. Gao, "An admission control scheme to provide end-to-end statistical QoS provision in IP networks," in *Proc. 21st IEEE Int. Conf. Perform., Comput., Commun.*, 2002, pp. 399–403.
- [33] M. Fidler and V. Sander, "A parameter based admission control for differentiated services networks," *J. Comput. Netw.*, vol. 44, no. 4, pp. 463–479, 2004.



Toufik Ahmed received the M.S. and Ph.D. degrees in computer science from the University of Versailles, Versailles, France, in 2000 and 2003, respectively.

He was a Visiting Scientist at the School of Computer Science, University of Waterloo, Waterloo, ON, Canada, in 2002. He is currently an Associate Research Fellow at PRISM Laboratory, University of Versailles. His main research activities concern quality-of-service for multimedia wired and wireless networks.



Ahmed Mehaoua received the Ph.D. degree in computer science from the University of Versailles, Versailles, France, in 1997.

He is an Associate Professor in the Computer Science Department, University of Versailles. He is currently the Leader of the Multimedia and Mobile Networking Group, CNRS PRISM Laboratory. Prior to joining the University of Versailles, he was with the University of Cambridge, Cambridge, U.K., as an Associate Research Fellow in the Center for Communications Systems Research and with the Hitachi

Telecom R&D Laboratory. From 1995 to 1998, he was with the Computer Science Research Institute, Montreal, Canada, and was involved in various R&D projects in telecommunication system management and quality-of-service management in broadband ATM networks. His main research interests include digital video communications over wired and wireless ATM and IP networks. His research has been supported by grants from government and corporate sources (IST, CNRS, RNRT, THALES).



Youssef Iraqi received the B.Sc. degree in computer engineering, with high honors, from Mohamed V University, Morocco, in 1995, and the M.S. and Ph.D. degrees in computer science from the University of Montreal, Montreal, Canada, in 2000 and 2003, respectively.

He is currently a Research Assistant Professor at the School of Computer Science, University of Waterloo, Waterloo, ON, Canada. From 1996 to 1998, he was a Research Assistant at the Computer Science Research Institute, Montreal, Canada. His

research interests include network and distributed systems management, resource management in multimedia wired and wireless networks, and wireless ad hoc networking.



Raouf Boutaba (M'01) received the M.S. and Ph.D. degrees in computer science from the University Pierre Marie Curie, Paris, France, in 1990 and 1994, respectively.

He is currently an Associate Professor in the School of Computer Science, University of Waterloo, Waterloo, ON, Canada. He conducts research in the areas of network and distributed systems management and resource management in multimedia wired and wireless networks. In these areas, he has published more than 120 papers in refereed journals and conference proceedings.

Dr. Boutaba is the recipient of the Premier's Research Excellence Award, a fellow of the faculty of mathematics of the University of Waterloo, and a distinguished lecturer of the IEEE Computer Society. He is the Chairman of the IFIP Working Group on Networks and Distributed Systems, the Vice Chair of the IEEE Communications Society Technical Committee on Information Infrastructure, and the Chair of the IEEE Communications Society Committee on Standards. He is on the Advisory Editorial Board of the *Journal of Network and Systems Management*, the Editorial Board of the *KIKS/IEEE Journal of Communications and Networks*, and the *Journal of Computer Networks*.

Meet In the Middle Cross-Layer Adaptation for Audiovisual Content Delivery

Ismail Djama, *Member, IEEE*, Toufik Ahmed, *Member, IEEE*, Abdelhamid Nafaa, *Member, IEEE*, and Raouf Boutaba, *Senior Member, IEEE*

Abstract—This paper describes a new architecture and implementation of an adaptive streaming system (e.g., Television over IP, Video on Demand) based on cross-layer interactions. At the center of the proposed architecture is the Meet In the Middle concept involving both bottom-up and top-down cross layer interactions. Each streaming session is entirely controlled at the RTP layer where we maintain a rich context that centralizes the collection of i) instantaneous network conditions measured at the underlying layers (i.e.: link, network, and transport layers) and ii) user- and terminal-triggered events that impose new real-time QoS adaptation strategies. Thus, each active multimedia session is tied to a broad range of parameters, which enable it to coordinate the QoS adaptation throughout the protocol layers and thus eliminating the overhead and preventing counter-productiveness among separate mechanisms implemented at different layers. The MPEG-21 framework is used to provide a common support for implementing and managing the end-to-end QoS of audio/video streams. Performance evaluations using Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) objective video quality metrics show the benefits of using the proposed Meet In the Middle cross-layer design compared to traditional media delivery approaches.

Index Terms—Cross-layer adaptation, forward error correction, link-layer quality, MPEG-21 multimedia framework, QoS metrics, real-time streaming.

I. INTRODUCTION

THE IEEE 802.11 WLAN standard is being accepted and deployed in many different environments such as companies, universities, government institutions, and public places (airports, train stations, etc.). The achievement of WLANs is draining an unprecedented research interest that is translating into a tremendous commercial success. Its adoption is favored by the promises of the forthcoming 802.11n specifications topping 540 Mbit/s (raw throughput) with the help of MIMO technology. 802.11-based networks appear as a serious alternative for wired ethernet, paving the way for QoS-enabled and added-value service provisioning for ubiquitous users. Thanks to the rise of powerful video compression techniques, such as H.264 and MPEG-4, it is now possible to combine video, audio, and data

within the same signal and transmit it over packet-based wireless networks. All these advances will help the emergence of new powerful multimedia applications with limitless possibilities and business promises. QoS-enabled multimedia services deployment is actually the main challenge to overcome in order to render WLANs an integral part of Network Operator's commercial offerings, and generally more attractive for the particular use.

Given the recent progress in both video encoding standards and wireless networks capacity, various in-door and out-door WLANs network operators are now more and more concerned by their ability to provide multimedia services with sustained QoS guarantees, while at the same time supporting a large number of heterogeneous wireless terminals. Providing QoS guarantees is an imperative in developing viable business models, while serving a maximum number of heterogeneous terminals is an obvious economical goal. Although the heterogeneity of terminal capabilities constitute an important burden on streaming systems, the QoS continuity of communication remains the main issue when streaming media over WLANs. Video adaptation techniques are clearly required in such environments with inherently varying link's capacity and quality. To achieve maximum efficiency, these techniques require a higher level of cross-layer interactions. This allows to i) translate application-layer QoS requirements into lower layers performance metrics such as link capacity and ii) reflect at the application-layer the short-term network fluctuations, captured at lower layers, such as interferences. QoS guarantees are particularly difficult to maintain in wireless environments where packet loss depends on the unpredictable occurrence and frequency of interferences. Error control techniques [e.g., FEC (Forward Error Correction) and ARQ (Automatic Retransmission reQuest)] are required in such environments. In particular, FEC is commonly used for real-time applications due to its proven scalability for multicast communications and the strict delay requirements of media streams.

With respect to video adaptation, the new MPEG-21 multimedia framework defines several parts for facilitating Digital Item (DI) consumption. MPEG-21 is intended to provide a general framework for multimedia access, delivery and consumption in order to enable seamless interoperability and content adaptation in distributed multimedia systems. More specifically, MPEG-21 defines two fundamental concepts which are Digital Item (DI) and User. DI represents an abstraction of a multimedia object (pictures, video or audio clip) and User is an abstraction of entities which interact with DIs. The MPEG-21 standard is divided into different parts. Each one handles a different aspect of the framework. Digital Item Adaptation (DIA), which is the MPEG-21 part 7 [1], specifies a set of tools to perform

Manuscript received February 22, 2007; revised July 26, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Anna Hac.

I. Djama and T. Ahmed are with the CNRS-LaBRI Lab, University of Bordeaux, F-33405, Talence, France (e-mail: djama@labri.fr; tad@labri.fr).

A. Nafaa is with the School of Computer Science and Informatics, University College of Dublin, Belfield, Dublin 4, Ireland (e-mail: nafaa@ieee.org).

R. Boutaba is with the School of Computer Science, University of Waterloo, Waterloo, ON, N2L 3G1, Canada (e-mail: rboutaba@bcr.uwaterloo.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMM.2007.911243

video adaptation, which includes Bitstreams Syntax Description (BSD), Usage Environment Description (UED), Adaptation QoS, and Universal Constraints Description (UCD). UED includes descriptive information related to user characteristics (e.g., user information and preferences, usage history, presentation preferences, audio language preferences, subtitles language preferences), terminal capabilities (e.g., codec capabilities and display capabilities, audio output capability, storage characteristics, battery capacities), network characteristics, and natural environment characteristics (e.g., location and time). The network characteristics provided by UED describes both static and dynamics network aspects. The static aspects are gathered in the network capabilities description (e.g., guaranteed bandwidth, delay transmission, error detection and correction), while the network conditions inform about the network dynamic status (e.g., packet delay, jitter, packet loss rate). An important feature of MPEG-21 resides in its ability to offer a rich, customizable, and interactive interface to users that enable them to change, at runtime, certain UED's attributes. The main challenge in future streaming systems is to enforce in a timely manner the UED requirements throughout the QoS adaptations implemented at different OSI layers. Note that terminal capabilities may also vary over the course of a multimedia session to reflect, for example, a codec reconfiguration/switching after a serious drop in battery level.

The way how network conditions are characterized by applications is an important factor that determines the effectiveness of adaptive mechanisms and remedial actions in face of transient network outages. In designing channel-aware video streaming systems, two interrelated challenging issues should be tackled: the accuracy of the effect of channel fluctuations and the effectiveness of application adaptation. The former consists in getting a thorough insight into channel fluctuations and their manifestations at application level by gathering a maximum number of QoS performance metrics at different levels of the protocol stack. Using QoS metrics of different protocol layers would deliver complementary information useful for building a consistent view of current delivery conditions. For instance, physical layer and transport layer measurements describe different manifestations of network fluctuations, and more importantly they give indications regarding different time scales. The latter challenging issue concerns the way the adaptation mechanisms interpret and react to those network fluctuations. For instance, 802.11 networks employ an inherent link-layer retransmission-based (ARQ) error control technique along with adaptive physical-layer coding/modulation to face frequent degradations in link quality. When used alone, 802.11 integrated adaptation mechanisms cannot avoid video degradations. In fact, these mechanisms are rather designed to increase the transmission reliability by either decreasing the transmission rate (coding/modulation) or increasing the average delay per packet transmission (ARQ). In certain circumstances, link quality degradations may be tackled at application-level (error control) with much more efficiency in terms of channel utilization. Particularly, application-level error control techniques (such as FEC) can improve application responsiveness by using physical layer signal strength measurements feedback. Indeed, metrics such as Received Signal Strength Indicator

(RSSI) represent an efficient indicator to predict the increase of packet loss rate. RSSI can be used to trigger an adaptive FEC mechanism to face packet losses and prevent user-perceived quality degradations. However, when the link signal strength degrades below a certain level, it is more appropriate to use a stronger modulation and coding channel at the physical layer. It is worth noting that application-level techniques are not anymore efficient when a large number of corrupted link layer frames occur. Clearly, coordination is of utmost importance for adaptation mechanisms to be efficient.

Our ultimate goal is to design an efficient video streaming system that is media, user/terminal, and channel aware. Our approach is based on a cross-layer design involving interactions between the application and the underlying layers of the protocol stack and a better integration of video coding semantics/metadata with the network transport mechanisms. The standardized MPEG-21 digital item management framework offers an ideal context for cross-layer interactions between distinct streaming system layers. Hence it provides a better consideration to environmental and external factors such as end-users' perceived quality, overall UED descriptions, and channel characteristics reported by the underlying protocol layers. Our approach, called Meet-In-the-Middle (MIM) cross layer, attempts to conciliate both the bottom-up and the top-down approaches usually undertaken in cross-layer design. Using top-down awareness, application-level QoS requirements are interpreted and enforced at the lower layers using appropriate QoS mechanisms. Meanwhile, bottom-up awareness allows to delivering valuable information on the wireless channel conditions to readjust upper-layer QoS adaptation mechanisms.

Meet-In-the-Middle centralizes the treatment of information, resulting from multi-layer network measurements and instantaneous user- and terminal-triggered constraints, at the Real-Time Transport Protocol (RTP) layer. The latter protocol offers the ideal context to manage both high-level QoS requirements and lower layers network and link measurements. By maintaining a large context (parameters set) related to a given multimedia session, it is possible to handle several aspects of real-time communication such as media synchronization, adaptive application framing, rate adjustment, link layer adaptation, etc. The objective here is to integrate and control all audiovisual adaptation mechanisms in a coordinated manner to minimize possible redundancies and counter-productiveness between layers, and thus improving the overall system efficiency. In this paper, we introduce the MIM cross-layer approach as well as its underlying principles. We provide concrete deployment scenarios in terms of cross-layer interactions to validate the effectiveness of our proposal. We evaluate the performance of a simple MIM cross-layer architecture where user/terminal requirements are communicated at session initiation phase while application constraints and network conditions measurements are continuously collected from both transport layer and PHY/MAC layer. The validation is achieved through the implementation of a prototype that provides live IPTV and VoD services to WLAN's heterogeneous receivers.

The rest of this paper is organized as follows: Section II discusses related works on video streaming over IP networks with

a special focus on cross-layer design approaches. Section III describes our proposed Meet-In-the-Middle (MIM) cross-layer adaptation and deals with the main implementation issues. Section IV presents performances evaluation results. Section VI concludes the paper.

II. BACKGROUND AND RELATED WORKS

A. Transport Services for Media Streaming

TCP and UDP are the most deployed transport services for real-time media streaming applications, although other protocols such as SCTP [2] and DCCP [3] have recently gained interest due to the support of new functionalities (e.g., flow control and multi-homing). Over the years, several enhancements of TCP and UDP have been proposed to better accommodate multimedia applications requirements. UDP-lite [4] is a typical example of these enhancements. By relaxing certain UDP and TCP constraints and implementing more efficient application-level error control schemes, audiovisual applications can reduce delays and excessive bandwidth consumption entailed by frequent packet retransmission. The IETF Audio Video Transport (AVT) working group has also specified the RTP (Real-time Transport Protocol) [5]. The origin of RTP stems from the growing need for more flexibility in managing various aspects of multimedia streams transmission. RTP may be used over any transport layer protocol (such as UDP, TCP, SCTP, etc.) though many redundant functionalities may exist with certain transport protocols. A typical example is the sequencing functionality provided by both TCP and RTP. Therefore, RTP is most commonly implemented over UDP. RTP follows the principles of application level framing and integrated layer processing proposed by *Clark and Tennenhouse* [6]. That is, RTP is intended to be malleable to provide the information required by a particular application and will often be integrated into the application level rather than being implemented as a separate layer. On the other hand, the RTP Control Protocol (RTCP), is used to monitor the quality of service and to convey information about the participants in on-going multimedia sessions. RTP/RTCP protocols are currently the “*de-facto*” Internet standards for real-time transport of various multimedia contents. Their specification is accompanied by documents (payload formats) that describe the specific encoding of different media types (RTP profile). RTP profile defines the meaning of each RTP payload and header fields depending on the transported media. RTP is usually implemented on top of UDP which offers no congestion control mechanisms and is thus unaware of network conditions and unfair towards other traffic flows. RTP/UDP may lead to an unfair behavior with respect to existing TCP flows. Since today’s Internet traffic is dominated by TCP (Web and P2P traffic), it is crucial, at least from network operators perspective, that UDP traffic behaves in a TCP-friendly manner so as to ensure that TCP-based applications continue to receive acceptable quality of service. Otherwise such applications will keep entering in a congestion avoidance phase, causing performance collapse and serious drop in the network throughput.

The average throughput of TCP can be inferred from end-to-end performance measurements such as round-trip-time (RTT) and packet loss. This observation has led to the definition

of TCP-Friendly Rate Control (TFRC- RFC 3448) [7] that can be performed by the end-hosts to maintain their rate within a certain level that matches a rate of an equivalent TCP session having the same conditions (i.e., the same session lifetime). The integration of lower level information such as packet loss and RTT into application-level processing was the first step towards the cross-layer design.

B. Cross-Layer Optimization

Cross-layer design has particularly arisen in wireless communications as an important paradigm to optimize the scarce wireless bandwidth utilization. It investigates situations where different OSI layers may cooperate to improve the ability of applications to achieve certain objectives such as QoS guarantees, power saving, or customization according to user preferences, etc.

Cross-layer design may be achieved by either integrating functionalities of different layers in a single protocol or simply establishing tight cooperation between adjacent (or non adjacent) layers.

In the former case, replication of information and redundancy of functionalities are avoided. This allows reducing the overhead and provides the means, thanks to the availability of a broad range of operational parameters in a single protocol, to implement advanced QoS mechanisms. The latter and most prevalent cross-layer design approach argues for richer inter-layer interactions to achieve better reactivity to network fluctuations and other external factors. The cross-layer parameters exchanged between layers depend on what functionality is being implemented as well as the objectives and constraints specific to the application being considered. A classification of useful parameters exchanged between layers in wireless communications can be found in [8]. In the following, we give three main categories of such parameters:

- 1) Channel state information (CSI) including location information, terminal capabilities, signal strength, interference level, etc.
- 2) QoS related parameters including delay, throughput, bit error rate (BER), and packet error rate (PER) measurements. Those parameters may be tracked at different layers.
- 3) Traffic patterns as perceived by each layer, including data traffic characteristics, knowledge of the data rate (constant or variable), data burstiness, data fragmentation, packet size (maximum transfer unit), and information about queue size.

Typically, cross-layer design approaches in wireless communications try to use the inherent variability of the wireless channel reported by the physical layer to adjust upper layers behaviors [9]. In fact, system components such as medium access control (MAC) protocols, routing algorithms, transport protocols, and application layer mechanisms may benefit from a certain degree of awareness about time-varying channel characteristics. The other way around, upper layer QoS constraints and application requirements may be translated into protocol behaviors enforced by appropriate mechanisms at the lower layers. For instance, by using transport layer statistics at RTP level, it is possible to adjust the application rate, error control techniques, buffering strategies, and so on.

C. Cross-Layer Video Streaming Architectures

Most of the works on cross-layer optimization have focused on MAC and PHY layers interactions in wireless environments. Very few works have considered higher level interactions such as the translation of user/terminal/application-level QoS requirements into well-proportioned QoS mechanisms.

In [10], the authors formalize the cross-layer design problem, discuss its challenges, and present a classification of possible solutions. The paper illustrates the relationships between layers and the challenge to find the best configuration that optimizes different metrics at different layers. The paper introduces the concept of “*coopetition*” between wireless stations, where a judicious mixture of competition and cooperation is often advantageous in competitive environments. When applied to wireless multimedia systems, “*coopetition*” changes the passive behavior of stations to adapt their transmission strategies to match available wireless and power resources, by enabling them to proactively influence the wireless systems dynamics through resource and information exchange.

In [11], the paper describes the recent advances in network modeling, QoS mapping, and QoS adaptation. The authors present a general block diagram of end-to-end QoS for video delivery over wireless networks. However, the end-to-end QoS is considered as a network-centric practice rather than application-level centric. Similarly works in [12]–[14], and [15] adopt a cross-layer design approach where both physical and MAC layer knowledge are shared with higher layers.

In [16], different error control and adaptation mechanisms available in several layers are evaluated for robust video transmission. Based on this evaluation, a new adaptive cross-layer protection strategy is proposed to enhance the robustness and efficiency of scalable video transmission. In [17], the authors propose a set of cross-layer techniques for adaptive video streaming over wireless networks. Data packets at application layer are decomposed on equal size radio link protocol (RLP) and FEC codes are applied based on RLP packets. The paper proposes also a priority-based Automatic Retransmission reQuest (ARQ) for corrupted RLP at application layer. The work in [18] present a new cross-layer content delivery architecture that is capable of receiving information from the network and adaptively tune transport parameters (e.g., bit rates) and other QoS mechanisms according to the underlying network conditions. The paper describes a service-aware IP transport architecture composed of a dynamic content-level audiovisual object classification model; a reliable application-level framing protocol with fine-grained TCP-friendly rate control and adaptive unequal error protection; and a service-level QoS mapping/packet-tagging algorithm for seamless IP differentiated service delivery. The obtained performance results demonstrate that by breaking the isolation of the OSI layers and by injecting content-level semantic and service-level requirements within the transport protocols, one can provide a more efficient support for multimedia services streaming.

More recently, a significant number of R&D activities are dedicated to QoS-sensitive multimedia delivery using cross-layer interactions. A number of such activities are conducted in the scope of European IST FP projects such as BRAIN, MIND, DRiVE, EVERES, PHOENIX, 4MORE, and ENTHRON.

Both 4MORE (4G MC-CDMA Multiple-Antenna System on Chip for Radio Enhancements, IST-507039) [19] and PHOENIX (Jointly Optimizing Multimedia Transmission in IP- Based Wireless Networks, IST-001812) [20] projects address cross-layer integration issues for multimedia streaming architectures. PHOENIX aims to develop a scheme to let the application world (source coding, ciphering) and the transmission world (channel coding, link modulation and coding) to further interact using the IPv6 protocol (network world), in order to improve the performance of multimedia communications. NEWCOM (Network of Excellence on Wireless Communications, IST-507325) [21] aims at identifying gaps in European knowledge on cross-layer practices to ultimately prepare an action plan for filling those gaps. The ENTHRON (IST-507637) [22] project proposes an integrated management solution which covers an entire audio-visual service distribution chain, including content generation and protection, distribution across networks, and reception at user terminals by integrating a cross-layer mapping of the application level QoS requirements to network level QoS concepts.

Similarly, our work falls in the area of audiovisual content delivery optimization using a cross-layer approach. Our approach however involves both top-down and bottom-up cross-layer interactions. Also, our work is not restricted to multimedia streaming over wireless networks. Different environments characteristics (terminals, users, multimedia applications, and networks) may be accommodated thanks to MPEG-21 tools. Finally, although the focus in this paper is on FEC and content adaptation techniques that are used as a proof of concept, other adaptation mechanisms may be as well envisioned.

III. MEET IN THE MIDDLE CROSS-LAYER ADAPTATION

The ultimate goal of Meet In the Middle (MIM) cross-layer adaptation is to ensure a seamless translation of user/application level QoS requirements into network level QoS metrics. At the same time, it allows reflecting at the application-level important network measurements carried out by the underlying layers.

User and terminal capabilities along with application and content specifications are essential knowledge when translating upper-layer QoS requirements into effective streaming adaptation mechanisms. In this top-down translation, a continuous flow of information including user, terminal, and application specifications are collected and used as metadata encoded using the MPEG-21 standard. In the same time, RTP-level QoS monitoring data (e.g., loss rate, delay, and jitter) and QoS statistics collected from real-time link layer measurements (e.g., Signal strength), network layer (e.g., Diffserv and queue information), and transport layer (packet loss distribution)¹, follow a bottom-up approach in the protocol stack. Fig. 1 gives a horizontal view of the proposed cross-layer approach that integrates content, terminal, user, and network management into a seamless MIM cross-layer architecture. The network-level measurements are collected on a per multimedia-session basis and used as part of a single context tied to an individual RTP

¹In this paper, we use the term “network-level” measurements to refer to different QoS measurements performed at the RTP layer and below (transport, network, and link layers). In contrast, higher-level (user, terminal, and coding) configurations are referred to as “service level” configurations.

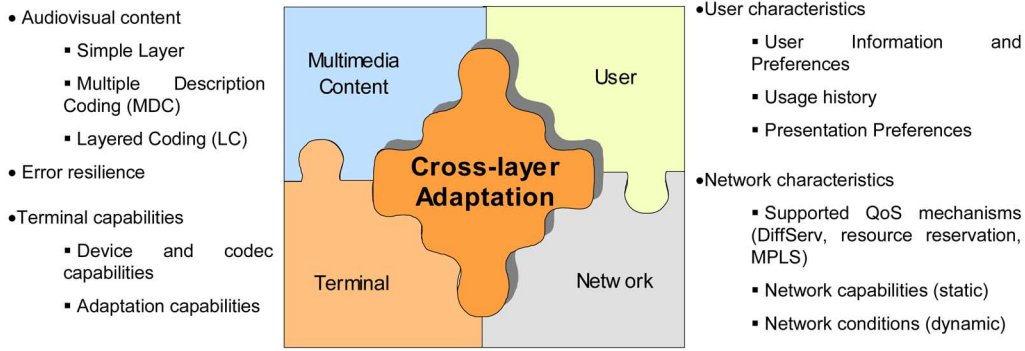


Fig. 1. Cross-layer adaptation based MPEG-21 metadata.

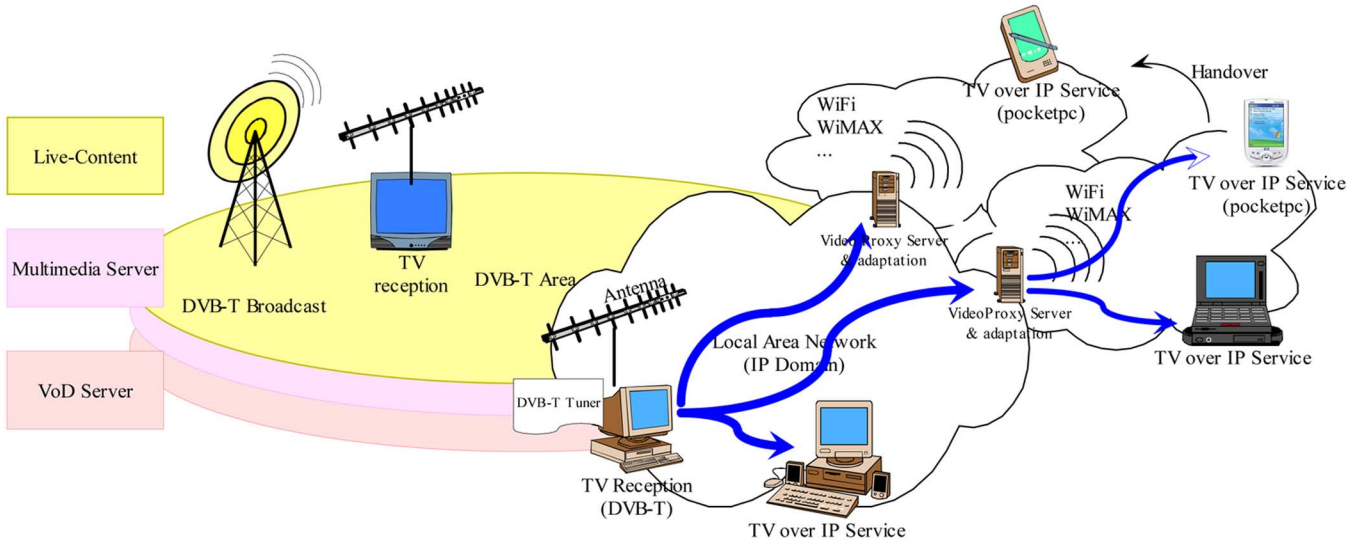


Fig. 2. Target architecture: Universal Media Access.

session (a context is maintained for each active multimedia stream). Thus, a different link-layer (respectively network, transport, application layer) adaptation strategy may take place for each handled multimedia stream regardless of possible QoS degradations experienced by other streams. For instance, link-layer transmission rate of a given stream may be deliberately decreased in response to harsh channel conditions due to receiver's distance from the access point (AP) or other environment-induced receiver's interferences. Based on the aforementioned QoS measurements, the server put into effect the appropriate combination of remedial actions available at different layers. Possible remedial actions include spatial/temporal resolution adaptation, per-pixel rate control, FEC redundancy adjustment, DiffServ re-marking, and MAC/PHY layer rate adaptation.

Live digital TV and VoD streaming services are used as a case study in our work, though MIM is not conceptually limited to this deployment scenario and may be, in practice, exploited with legacy Internet real-time and non real-time services. Our target architecture is described in Fig. 2. In this architecture, we particularly focus on WLAN's receivers. The WLAN is considered as a last-mile connection which is subject to quality degradations. We assume that multicasting multimedia streams between the various servers and TV receivers in the core and the wired po-

tions of the network don't suffer from any degradation thanks to large capacities and resources over-provisioned. Quality degradations are essentially provoked by the wireless access network.

Signaling and "horizontal" interactions between receivers and servers may be supported by using several different signaling protocols (e.g., RTCP, RTSP, HTTP, and MMS). These signaling protocols allow receivers to report to the servers certain i) QoS metrics measurements achieved at different layers and ii) higher-level events triggered by terminal- and user- capabilities/preferences changes. We have extended the RTCP report to carry additional network-level QoS information such as loss distribution pattern, loss rate before, and after FEC recovery. Besides, we use RTSP together to transmit MPEG-21-compliant metadata regarding user, terminal, and encoding capabilities from the end-user's terminal to the streaming server. RTSP may be also used with an already active session to report on new events triggered by the user or the terminal. This protocol uses the IETF common format to express media and session descriptions, namely the session description protocol (SDP) [23]. Note that from the receivers' perspective, the content is exclusively located at the MIM-enabled Server (i.e., Video Proxy Server and Adaptation) as shown in Fig. 2. In the rest of this paper, we use the term video server to refer to the Video Proxy Server and Adaptation.

A. User/Application Level QoS Description Using MPEG-21 Metadata

SDP is basically used to describe i) session parameters for announcements or invitations and ii) the capabilities of a system, and possibly for providing a choice between a numbers of alternatives. In other words, SDP is used to convey information related to the multimedia stream before the session activation. This may include information related to both transport configuration (RTP and RTCP port number, RTP Payload Type definition, RTP payload header signalization, etc.) and multimedia encoding configuration (encoder parameters).

Current multimedia applications have higher demands in terms of advanced features and customizable configurations. However, SDP falls short in coping with these new demands. For instance, SDP does not provide information related to user profile (user information and preferences, usage history, presentation preferences, etc.), neither does it support the description of terminal capabilities such as adaptation capabilities. QoS parameters for different protocols such as traffic specification and flow specification or DSCP (Diffserv Code Point) for IP QoS differentiation are not supported by SDP. Yet, these QoS parameters need to be specified somehow (e.g., using out of band signaling). These deficiencies led the IETF MMUSIC working group to investigate a new generation of the SDP protocol namely SDPng (SDP next generation) to support various long-term extensions. On the other hand, MPEG-21 can fill the SDP gap by providing metadata related to the user, terminal, content and network characteristics. We believe that building SDP on top of MPEG-21 will provide the means to enable access to multimedia content under a wide range of delivery conditions and usage environments. In this perspective The IETF draft [24] presents a practical approach for harmonizing MPEG-21 with SDPng. In our work, SDPng is used to specify the context-layer of user preferences, content capabilities, and terminal requirements at session initiation stage. The specified parameters are then taken into account at the server level to appropriately configure the content and other adaptation mechanisms.

B. Cross-Layer Adaptation

Fig. 3 depicts the MIM cross-layer architecture and highlights the main top-down and bottom-up interactions between the involved layers. More precisely, the figure depicts the network architecture of a server which, unlike the receiver's, includes the MIM cross-layer adaptation engine. The latter is the core component centralizing i) server perception of current network-level conditions, ii) service-level configurations, and iii) current reported events and feedbacks from the receiver (i.e., client).

Detailed format of user characteristics, terminal capabilities, and content specifications are described in MPEG-21 metadata based on XML. The XML-compliant metadata coding constitutes the vehicle for vertical communications that take place between adjacent layers at each end-system. For example, and without loss of generality, a terminal that supports 8-bit color scheme could express its requirements down-to the RTP through an XML-based MPEG-21 element as follows (`< m21-dia:ColorBitDepth blue="8" green="8" red="8"/>`). Besides, the frame rate supported by a terminal could be expressed under the attribute *"refreshRate"* while the display capability of its screen could be

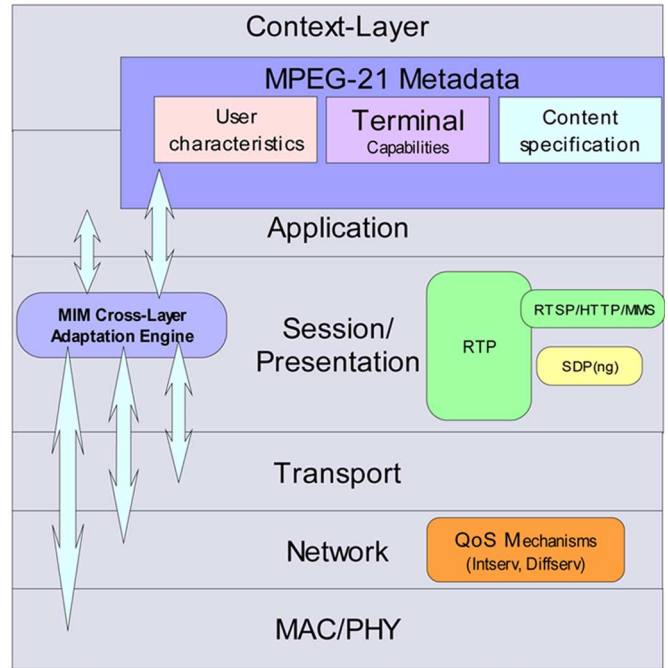


Fig. 3. MIM cross-layer interaction.

expressed in terms of pixels under the two attributes *"horizontal"* and *"vertical."* The transmission of terminal capabilities across the layers is an example of vertical communications that take place using MIM. Similarly, bottom-up interactions are used to communicate link and network performance measurements to upper layers as we describe it later in this paper.

XML-based requirements are enclosed into an SDPng message and then exchanged between end-systems (server and receiver) through the appropriate streaming session signaling (respectively maintenance) protocol such as RTSP. At the server, the SDPng message is parsed to extract the terminal (respectively user and content) capabilities that are again passed to RTP using the same MPEG-21 based XML encoding. Note that not only the receiver sends its capabilities to the server, but the server may as well send to the client its coding or transport capabilities at the streaming session initiation stage. This server to receiver communication may include video configuration specifics, streaming session IP address and port numbers, FEC stream decoding parameters, possible video rate adaptation granularity, and possible video temporal/spatial resolutions.

RTCP-based feedbacks reporting is another type of multimedia session signaling communication that is rather frequent and continuous during the session life-time. Typical information conveyed within RTCP reports are usual QoS metrics performance as well as certain advanced measurements such as loss pattern measurements. At the server side, this QoS feedback information is analyzed at the RTP level and then reflected at different layers by leveraging appropriate QoS adaptation mechanisms ranging from video encoding adaptation to PHY/MAC encoding and modulation re-setting. Note that not only network conditions are continuously reported, but also other user-centric parameters of UED (such as screen format, battery depletion level, user preferences, etc.) may be as well passed and continuously updated using RTSP.

At the network layer, an advanced interaction with upper and lower layers is necessary to ensure QoS continuity. A QoS matching and packet tagging algorithm is used to map application level QoS to network level QoS, and thus ensuring seamless IP differentiated service delivery. For instance, let's take the example of a video content structured as a single layer stream with I, P, and B frames. Each video frame could be carried over the IP Diffserv network with a particular DSCP code point. From the perspective of video distortion, it is known that an I-frame is more important than a P-frame which, in turn, is more important than a B-frame. Consequently, IP packets carrying I-frame fragments (or slices) will be marked with low drop precedence compared to IP packets carrying P-frame fragments, and so on.

Dynamic changing of context-related information should be also carefully considered (dynamic UED). Conventional advanced streaming systems collect network characteristics at the session initiation stage to statically configure basic video encoding parameters that usually remain unchanged throughout the streaming session lifetime. Using MIM, it is possible for the user to clearly express certain MPEG-21 based "*NetworkCharacteristic*" and thus put restrictions on maximum bandwidth ("*maxCapacity*"), average throughput, supported QoS classes, etc. These elements are also considered for dynamic adaptation. Some of them can be carried by end-to-end protocols such as RTCP. Several adaptation mechanisms can be performed at the server; some of which are described in the next section. It is worth noting that the main goal of this paper is not to focus on some adaptation scenarios or techniques but is rather to discuss and analyze the benefits of cross-layer adaptation in regards to collaboration between layers from MIM perspective.

C. Cross-layer Adaptation Strategies

This section describes three different cross-layer adaptation mechanisms: Link layer rate adaptation; adaptive FEC redundancy for video streams; and video content adaptation (temporal/spatial resolution and SNR level). All three mechanisms are based on MIM cross-layer interactions. We believe that the combination of these three adaptation mechanisms is sufficient to assess the ability of MIM cross-layer adaptation in facing changes in network conditions while meeting users' QoS expectations.

1) *Link Layer Rate Adaptation*: The 802.11 physical layer specifications introduce multirate capabilities which allow the physical layer to provide different channel coding and modulation. The link rate is strongly related to these two parameters. In fact, complex channel coding and modulation produce a higher link rates but are more sensitive to noise. On the other hand, lower link rates are based on stronger channel coding and modulation which are more resilient to noise. Therefore, it is more suitable to use a higher link rate when the wireless station are close to the Access Point and progressively decrease this link rate as the station moves away. The link layer rate adaptation automatically selects a transmission rate from a set of allowed rates based on the transmission conditions of the last frames. This automatic selection is performed by the Rate Control Algorithm (RCA) at Link layer. In [25], the authors identify three classes of RCAs: Statistics-based RCAs, SNR-based RCAs, and Hybrid RCAs. The Statistics-based RCAs maintain statistics information about the transmission conditions and the

achievable throughput to adapt the rate consequently. Authors in [26], studied the performance of Auto Rate Fallback (ARF) algorithm, which belongs to statistics-based RCAs category since it is based on a continuous monitoring of the number of successfully acknowledged messages at link layer. The SNR-based RCA adapts the rate directly according to the SNR perceived by receivers since the appropriate channel coding and modulation are related to signal quality (SNR level). The link rate adaptation strategies proposed in [27] and [28] are based on these RCAs range. The third RCAs type, namely Hybrid RCAs, use a combination between the two above-mentioned techniques to take advantages of each of them and minimize their shortcomings.

The rate control algorithms (RCA) are used to allow maximizing the data throughput while minimizing the packet loss by switching between different rates (e.g., from 54 Mbit/s to 11 Mbit/s). In this context, work in [29] analyzes the streaming video quality and captures wireless LAN characteristics across network and wireless link layers. Authors investigate possible WLAN performance indicators that may be used to predict the streaming video quality. The results show that the wireless RSSI and average wireless link capacity are the most accurate indicators to predict the performance of streaming video over wireless LANs. In the similar optic, we have shown in [30] that when the signal strength decreases, the MAC frame error rate increases consequently. The "Sample" RCA, which can be classified in statistics-based RCAs, uses a transmission time comparison achieved by different link rates. We demonstrated that adapting the video content according to link rate, decided by RCA, improves significantly the overall perceived video quality.

Based on this concept, it is clear that the correlation between the data rate and signal strength has to be taken into consideration in our MIM cross-layer adaptation where the signal quality measurements are used as input for triggering different adaptations.

2) *Forward Error Correction (FEC) Adaptation*: Packet loss is a problem that considerably affects the quality of received video quality at the client. It may lead to a very destructive effect on the reconstructed video sequence, because of video frames dependencies. This phenomenon is usually referred to as "*error propagation*." Packet loss may happen at different levels and due to different reasons. In wired packet-switched networks, congestion is the first cause of packet loss. Entire packets can be discarded by routers. Whereas in wireless networks, the transmission channel may cause frequent bit errors. Corrupted packets are discarded and hence considered lost.

Forward error correction can be applied at many levels from bit level up to packet level. In a bit level FEC, a bit is considered as a symbol while in packet level FEC, a symbol is a packet. Packet level FEC consists of producing "*k*" redundant packets from "*k*" original ones. FEC packet is generally based on erasure coding and its usefulness lies on that i) a single parity packet can be used to correct different single-packet losses in a group of packets, i.e., packets belonging to a given FEC block; ii) bit level FEC is unable to recover a completely lost or delayed packet; and iii) when using a bit level FEC, a corrupted packet is already detected and discarded at link layer (respectively transport layer) before being available at application level; hence, using a bit-level FEC at application level implies disabling error detection mechanisms (CRC and Checksum) of underlying layers.

Even though most of existing wireless access networks use adaptive coding and modulation schemes integrated to the link layer (see Section III-C.1), packet-level FEC protocols are usually required. As shown in [31], wireless communication experiences i) fast fading and white Gaussian noise, which are addressed by the integrated physical layer coding and ii) slow fading (e.g., when entering a tunnel), which is addressed by a packet level FEC encoding. These two levels of FEC encoding are complementary, each one addressing different problems. Clearly, there is a need for packet-level FEC protection, in addition to bit level FEC, to increase wireless multimedia communications reliability.

Typical packet-level FEC protocol that uses k media packets to produce n packets, among which $h = n - k$ parity packets, have the capacity to overcome up to h packet loss (when using MDS codes). This basically provides a resiliency against a maximum packet loss rate of $p = h/n$ when considering that even FEC packets may be affected by loss. Thus, based on the average packet loss rate measurements, such as those provided by the RTCP feedback, it is possible to adjust the level of redundancy each time as follows:

$$h = \frac{p \cdot k}{(1 - p)}. \quad (1)$$

Based on the number (k) of media packets to protect and the measured mean loss rate (p), the number of FEC packets can be easily determined. Note that at the receiver, the loss rate (p) is measured at the transport level, which means before applying FEC recovery. This way, the server gets a consistent picture about the current network conditions in order to adjust the FEC redundancy. The overhead introduced by the FEC redundancy should be tackled by content adaptation mechanism such as *transrating* as proposed in this paper to maintain smooth bandwidth utilization.

Besides the transport-level monitoring and the network status measurements that may be carried out at different levels, delivering different indications are of great importance for FEC adaptation mechanisms. PHY/MAC layers signal strength measurements are among the most important QoS metrics in statistically shared environments like WLANs. Further, other advanced measurements such as loss pattern (packet loss distribution) may be useful for adjusting FEC transmission [32].

The frequency with which the network loss rate is reported to the sender may deteriorate the responsiveness of FEC schemes leading to suboptimal FEC efficiency. A high frequency would enhance the responsiveness at the sender while causing high variations between successive measurements (e.g., leading to instability of the system), not to mention the uncured excessive feedback overhead. In turn, a low frequency would trade-off good stability and low overhead for poor reactivity. In our case, we use RTCP reporting with a fixed frequency (up to 5% of RTP session bandwidth as recommended in the IETF standard). Furthermore, the RSSI is measured for each connected client at the access point which plays the role of proxy and video adaptation gateway. Clearly, using RSSI measurements, the adaptation gateway will have access to a coherent and up-to-date view of the network conditions perceived by each receiver. Indeed, the time-scale of physical layer measurement is very small and

revealing of the short-term network conditions. This would considerably improve the responsiveness of FEC redundancy control as it takes into consideration the short-term degradation of signal quality. The MIM cross-layer FEC redundancy is based on both signal strength measurement at the client side and network packet loss ratio. It is adjusted dynamically to overcome as much as possible the video quality degradation. With the help of the content adaptation, the amount of bandwidth used can be maintained while FEC redundancy is added.

3) *Content Adaptation*: Transmitting packet video streams over WLAN encounters the problem of network capacity variation (i.e., bandwidth fluctuation) as the signal strength is unpredictable. The bandwidth of the path between the sender and the receiver is the most important characteristic that directly affects the quality of video services. It is however generally time varying and hardly predictable. If the sender transmits more than the available bandwidth, video packets may be lost or may experience excessive delays. A common technique to deal with bandwidth variation is to use adaptive content streaming. The server estimates the available bandwidth and then adapts its sending rate to match the available bandwidth. This technique is widely dependent on the video coding flexibility and features allowing for example video rate adaptation, multi-resolution streams adjustment, etc. The Scalable Video Coding (SVC) [33] defines three-dimensional scalability to allow for adaptation in heterogeneous environment by simply truncating appropriate bit streams parts. Temporal, spatial, and SNR (Signal to Noise Ratio) scalability are among the well known techniques used to tackle bandwidth variation and fluctuation. The temporal scalability is based on video frame rate, in which a higher quality layer corresponds to a higher video frame rate. In the spatial scalability, the quality layer has different video frame size and the quality increases by increasing the frame size of video. Finally, the SNR scalability is based on quantification factor which represents the visual quality of video pictures. These three dimensions are considered in our MIM cross-layer adaptation. Initially, content adaptation is performed at multimedia session initiation phase based on user preference (respectively terminal capabilities and encoding constraints) carried out using MPEG-21 UED; it is possible to select a streaming format out of various temporal/spatial resolution and SNR levels.

During multimedia transmission phase, we keep unchanged the temporal/spatial resolution for transmitted video. However, the SNR level is adapted by the server to control the video throughput. In fact, according to the cross-layer parameter (signal strength quality, change in link rate, and packet loss ratio) the server adjusts the FEC redundancy and adapts its sending rate by *transrating* the video content using quantification factor to overcome the overhead introduced by the FEC.

D. Implementation

The MIM cross-layer adaptation architecture is depicted in Fig. 4. It is composed of a media server and a client. The server streams audiovisual content to the client via an IP network using the RTP protocol.

Consider the use case of VoD streaming over IP with a multitude of users and 4 devices. At the connection phase, the client requests the content through the RTSP protocol. The

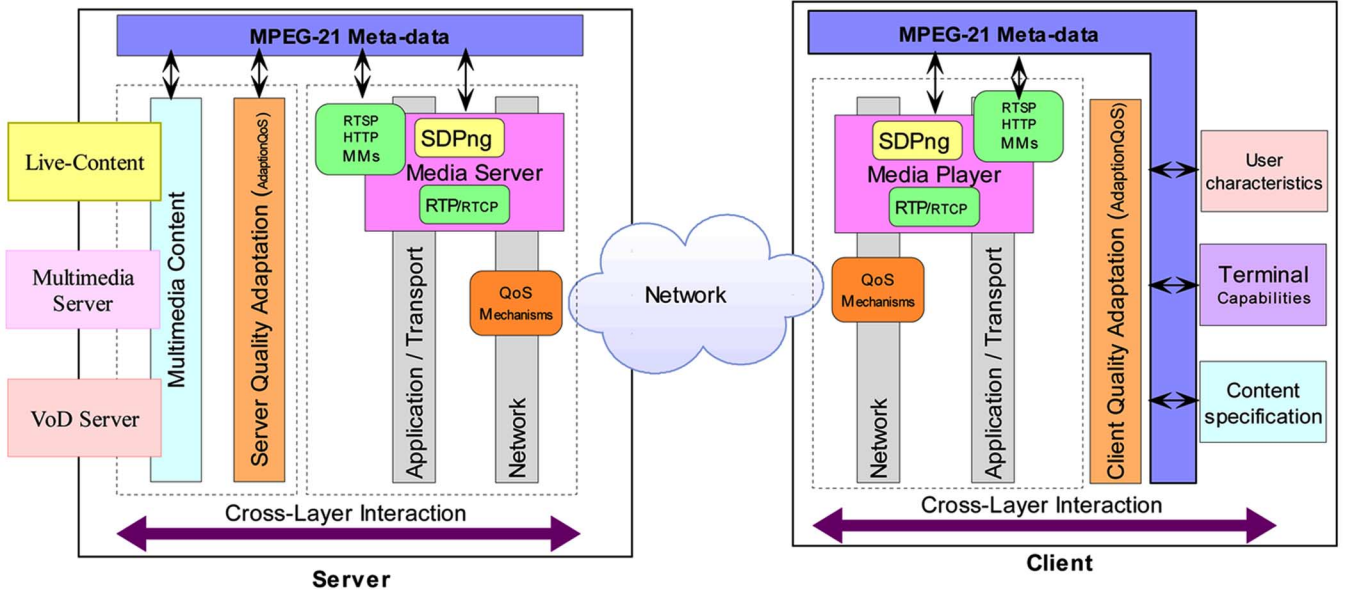


Fig. 4. MPEG-21-enabled cross-layer adaptation architecture.

MPEG-21 descriptions of UED are encoded in SDPng format and enclosed in HTTP or RTSP messages. The server's quality adaptation engine adapts the content based on UED descriptors (color scheme, refresh rate, screen size, codecs capabilities, and network characteristics, etc.). This represents the first MIM cross-layer interaction from the context-layer down-to the RTP layer. The server adapts dynamically to the changing conditions using the client signal strength, physical link rate, and packet loss-ratio that are carried through RTCP feedback. At this stage, the server adjusts the application-level QoS mechanisms such as FEC redundancy amount and video throughput.

During the streaming session lifetime, different types of QoS measurements can be collected and used to refresh UED descriptors made available at the server side for adaptation purposes. Such variety of measurements is not supported in our current implementation which is only limited to the aforementioned cross-layer parameters. In fact, standard RTCP "fraction lost" before and after FEC correction, "cumulative number of packets lost," and "inter-arrival jitter" allow the server to be continuously aware of long- and short-term changes in network conditions.

In our implementation, both the client and the server are based on open source project VideoLan (VLC) [34]. VLC is a highly portable multimedia server and player for various audio and video formats (MPEG-1, MPEG-2, MPEG-4, DivX, mp3, ogg ...) as well as DVDs, VCDs, DVB-S/T and various streaming protocols. It is used as a server to stream in unicast or multicast in IPv4 and IPv6 networks. VLC protocols have been extended to meet the requirements of our adaptation system. Our VoD server is based on VLC-0.8.5 running Linux 2.6.17 kernel.

IV. PERFORMANCE EVALUATION

We have deployed a test-bed to experiment with our proposed MIM cross-layer adaptation and evaluate its performance. The performance evaluation takes into account several aspects entailed by a practical deployment of the video streaming system,

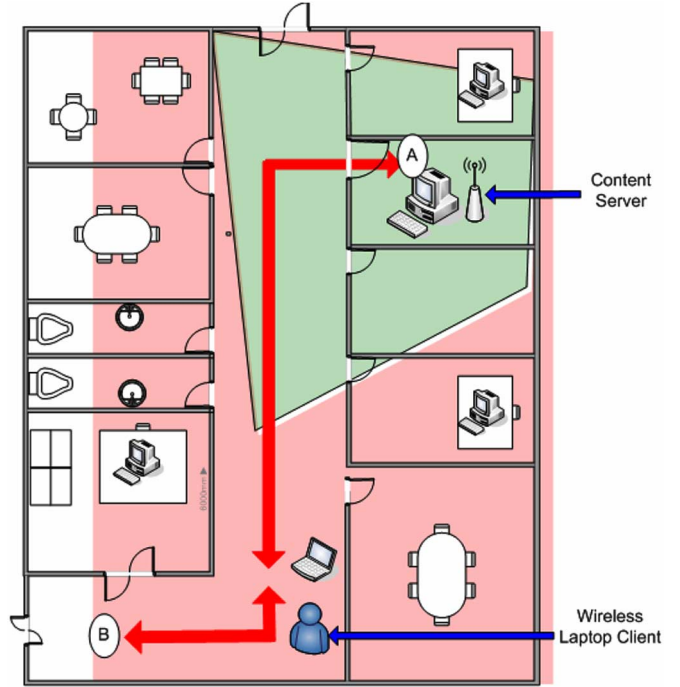


Fig. 5. Test-bed configuration and RSSI quality map of the environment.

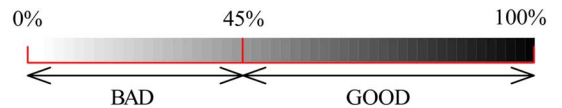


Fig. 6. RSSI percentage used in MIM.

with a special focus on the appropriate QoS performance metrics to be measured by the network operator. The configuration of the test-bed is illustrated in Fig. 5.

In our experiments, we used an MPEG-4 coded Akiyo video sequence stream as a reference for testing. The Akiyo video sequence is 300-frames length with a video frame rate of 25

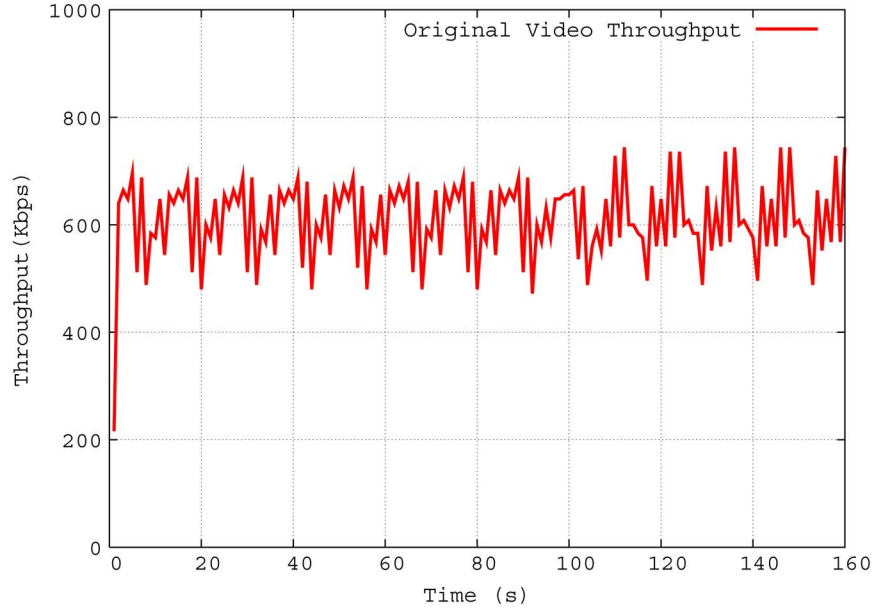


Fig. 7. Instantaneous throughput of the reference video.

frames per second. The server streams the video sequence continuously in a loop which we believe is fairly sufficient to highlight the main benefits of our architecture. The experiment duration is about 160 s during which the wireless mobile user moves from location A (content server) to location B, and then comes back to location A with a regular pace as illustrated by the red arrow in Fig. 5. The test-bed is deployed in a closed environment (laboratory), and the mobile user may be separated by up to three walls when it is at maximum distance from the Access Point. Both Access point and client wireless card are based on Atheros 802.11a/b/g chipsets. The Atheros chipset provides (through the MadWiFi open source driver [35]) two parameters to report on the channel state. The first is called “*signal strength*” which is measured in *dBm* with an interval range of $[-96, 0]$, the higher strength is 0 dBm and the lower is -96 dBm. The second parameter is called “*RSSI*” and provides a signal quality. The RSSI does not have unit and their range interval is $[0, 94]$ where 0 represents the worst signal quality and 94 the best one. Obviously, there is a strong correlation between these two parameters variations as they both report different metrics of the link quality. However, for our MIM cross-layer adaptation, we choose to use the RSSI percentage by dividing the instantaneous RSSI value by the maximum RSSI which corresponds to index 94. The RSSI percentage provides a non-relative value which is more appropriate to measure the signal quality variation with acceptable accuracy and granularity. Fig. 6 shows the RSSI percentage used to steer the adaptation according to two ranges: 1) $[0\%–45\%]$ is considered to represent bad link conditions and 2) $[45\%–100\%]$ is considered to represent good link conditions. We also depict in Fig. 5 a rough estimation of the signal quality map (Good/Bad) measured in our lab environment. The locations where the signal quality is good are represented by green color, while red color represents locations where the signal quality is below the threshold (45%).

During the experiment, the same mobility pattern is each time assumed where the user moves from A to B. Due to time-

Packet loss rate	No loss	No FEC No overhead
	<5%	FEC(10,8) → Overhead +25%
	>5%	FEC(10,7) → Overhead +42%

Fig. 8. Example of simple adaptation strategies for scenario 2.

varying wireless channel conditions that depends on several unpredictable phenomenon and interferences, each run of the experiment produced slightly different results, but with always the same trend and within an acceptable disparity margin. In other words, the disparity in the measured signal quality is fairly minor. Initially, at time $t = 0$ the mobile user moves from A to B with a speed of about 1 meter per second. At time $t = 60$ seconds, the user arrives at location B. The user stays in location B until time $t = 100$ seconds, where he experiences the worst channel condition. Afterward, the user returns back to location A and reaches its initial position at time $t = 160$ seconds.

The Akiyo video sequence used in the experiment has a size of 368×242 pixels, an average throughput of 609 Kbps, and a peak rate of 760 Kbps as shown in Fig. 7. No audio traffic is sent during the experiment. In the rest of this section, we refer to the Akiyo video traffic as “*reference sequence*,” which is available at [36]. In order to appropriately receive and display the video streams, we use a wireless laptop with 1.5 GHz processing capabilities and running a modified version of VLC-0.8.5. Video quality measurements are performed using objective metrics (PSNR and SSIM) that show the user-perceived quality gain entailed by the use of MIM cross-layer adaptation.

In the following, we do not consider the static adaptation carried by SDPng and performed at the connection phase. We focus only on different adaptation scenarios performed during the course of the streaming session. Although both adaptations have

RSSI level	Packet loss rate			
	Good [45% - 100%]	No loss	<5%	>5%
		- No FEC - No transrating	- FEC(10,8) → Overhead +25% - Adapt the video → Transrating the video -25%	- FEC(10,7) → Overhead +42% - Adapt the video → Transrating the video -42%
	Bad [0%-45%]	- FEC(10,8) → Overhead +25% - Adapt the video → Transrating the video -25%	- FEC(10,7) → Overhead +42% - Adapt the video → Transrating the video -42%	- FEC(10,7) → Overhead +42% - Adapt the video → Transrating the video -42%

Fig. 9. Example of simple adaptation strategies for scenario 3.

equal importance, we have focused on the dynamic adaptation for its ability to counter unforeseen events such as fluctuations in network conditions. With MIM cross-layer adaptation, the receiver continuously measures and reports the instantaneous packet loss ratio using RTCP. The server embedded at the access point measures the signal quality of all currently active users using the API of MadWiFi open source driver [35]. The above measurements (i.e., packet loss ratio, signal quality) are used by the server to re-adjust packet-level FEC redundancy transmission and to perform content adaptation. While the FEC increases the overhead, the video rate adaptation aims to reduce the video throughput. Both mechanisms applied together allow maintaining a smooth video bandwidth at the server. Note that the packet loss rates measures are collected at the receiver side before and after FEC recovery. This allows to accurately reporting on real network conditions experienced by the video traffic.

For the sake of comparison, we have tested and evaluated the performance of three streaming application scenarios, namely: a conventional streaming system, a streaming system with adaptive FEC only, and the MIM cross-layer adaptation. The evaluation focus on assessing the performance gain that may result from the combination of multi-layer QoS measurements such as transport-level loss rate and RSSI at PHY/MAC layer. The three tested scenarios are explained in the following.

- **Scenario 1:** reference sequence streaming with a conventional streaming system.
- **Scenario 2:** reference sequence streaming with adaptive streaming system. The server uses RTCP's packet loss rate measurements to adjust its FEC redundancy transmission. PHY/MAC-layer measurements are not collected. The adaptation strategies are shown in Fig. 8.
- **Scenario 3:** reference sequence streaming with the MIM cross-layer adaptation. We perform two different measurements to adapt the FEC redundancy: signal quality at link layer and packet loss ratio measurements at transport-level. The adaptation strategies are shown in Fig. 9.

The adaptation strategies introduced above are rather simple as they are intended to evaluate the benefits of the MIM cross-layer adaptation and to gain insight into the problem of coordinating different measurements at different layers. More specifically, we intend to show the advantage of using short-term and

long-term performance measurements at different layers of the protocol stack such as signal quality and RTCP packet loss ratio. The RTCP loss threshold is fixed to 5% in the adaptation strategies in an arbitrary way so as to show the MIM cross-layer adaptation in our experimental test-bed. The objective is to capture the visual degradations (i.e., the subjective quality evaluation) entailed by packet loss and to emphasize the gain in performances achieved by MIM-based adaptation strategies. In real system, this threshold can be fixed by service provider according to the service level subscribed to by an end-user or according to service characteristics (loss ratio tolerated by content delivered to the client). More specifically, the loss threshold can be chosen directly by an end-user and integrated to the Usage Environment Description as part of the user preferences which are transmitted to the server during RTSP negotiation. The threshold (e.g., max loss rate) that determines the aggressiveness of MIM adaptation in response to network degradations is also dependent on the streamed media and their resiliency to packet loss. While it is commonly accepted that the video streams are, to certain extent, fairly resilient to packet loss, it is clear that different video encoding formats may present different resiliency levels. Further, it is also quite obvious that even different encoding efficiencies (rates, SNR level) with the same format results in different loss resiliencies. The more efficient the video encoding is, the more sensitive to error propagation effect, and the less the resiliency to loss the playback video is. The issue of establishing the optimal adaptation rules with the most appropriate adaptation thresholds is a multidimensional problem that should be addressed by taking into accounts both the user-to-service provider contracted SLA, the loss resiliency of the streamed media, and more generally the level of acceptance of users in terms of perceived video quality.

In order to deal with short term oscillations in signal quality measurement in scenario 3, we used a low-pass filter to smooth the consecutive measured value. The used low-pass filter is an Exponential Weighted Moving Average (EWMA) able to quickly detect unusual situations. It uses one exponential smoothing parameter to give more weight to recent observations and less weight to older observations or vice-versa, as shown in (2)

$$WRSSI \leftarrow (1 - \lambda) * RSSI + \lambda * WRSSI. \quad (2)$$

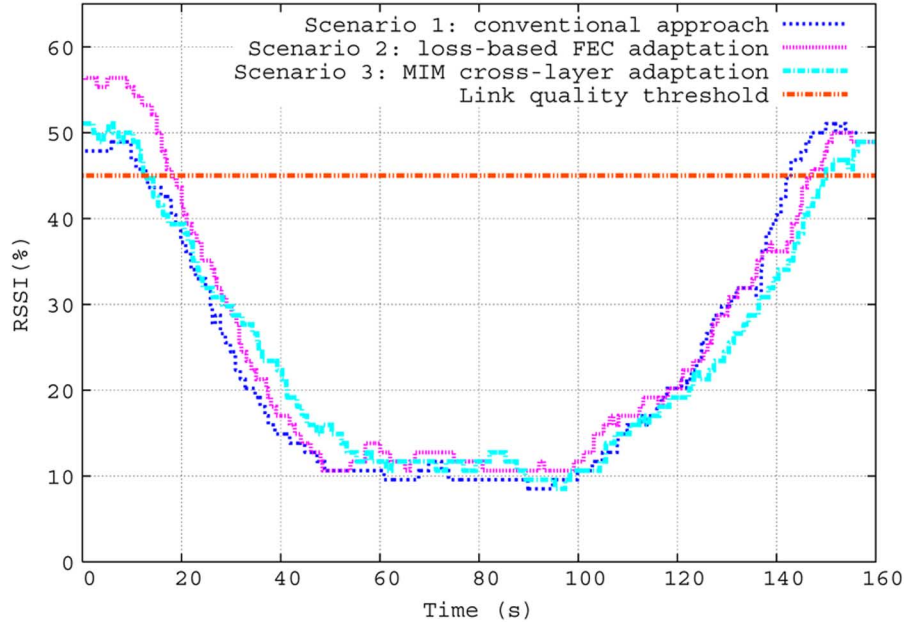


Fig. 10. Instantaneous smoothed measurement of Signal quality(%).

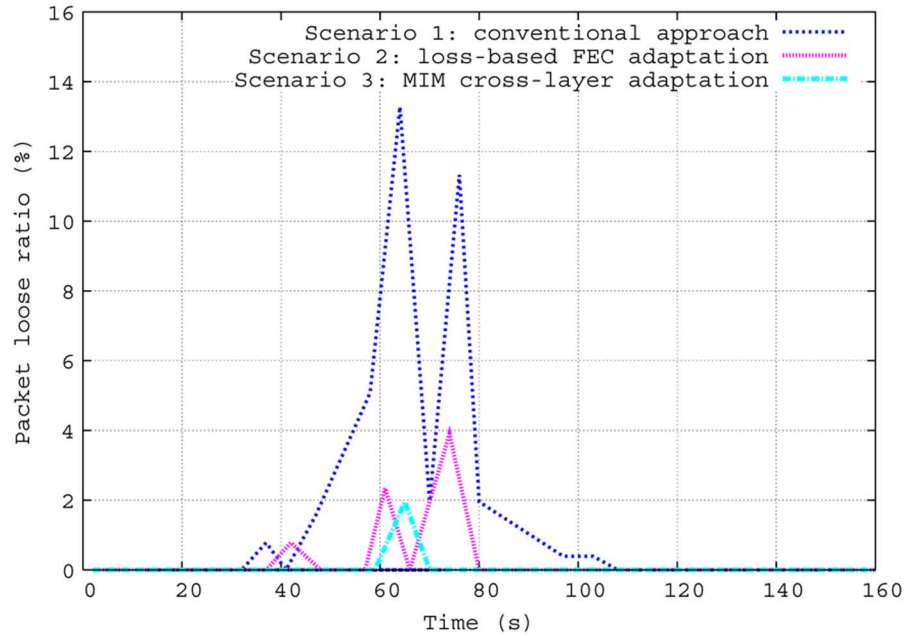


Fig. 11. Instantaneous packet loss ratio.

During the experiments, we choose $\lambda = 0.8$ to detect small shifts in link quality while limiting the effect of transient fluctuations. The adaptation decision is then based on the smoothed value of RSSI, namely WRSSI, rather than on instantaneous measurements.

Fig. 10 shows the measured signal quality experienced by the mobile user while roaming. The measurements are performed using Atheros driver which gives RSSI values ranging from 0 (0%) to 94 (100%) of quality percentage. As mentioned earlier, the signal quality threshold is set to 45%. In all three tested scenarios the experienced signal quality is more or less the same as the user moves, though there is some measurement variability

between the different scenarios due to unpredictable noises and interferences which cannot be avoided when dealing with real experiments and test-beds. The signal quality drops significantly when the user is 30 meters away from location A with a RSSI of 10% as the average received signal quality. This leads to significant packet loss as shown in Fig. 11 for Scenarios 1 and 2. Here, packet loss ratios are measured after FEC recovery when FEC is used (Scenario 2 and 3). As expected, the measured packet loss ratio increases proportionally with the drop in the signal quality observed in Scenarios 1, 2, and 3 between $t = 40$ s and $t = 100$ s. The measured loss ratio for Scenario 1 can serve as a reference to assess the network-induced losses since no FEC recovering is

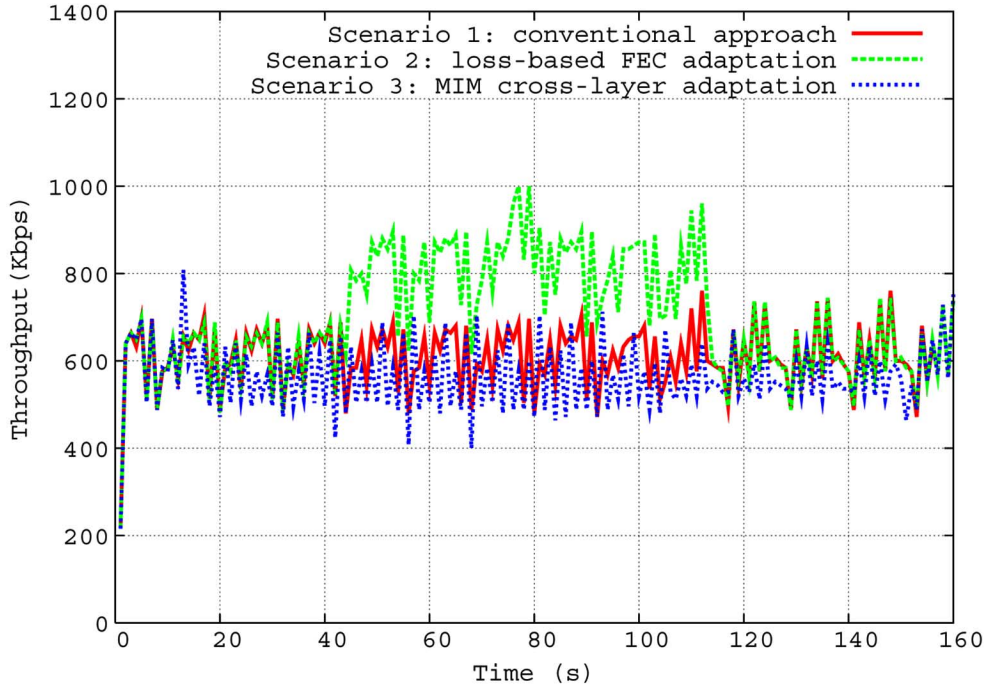


Fig. 12. Instantaneous video throughput achieved at the server side.

applied in this scenario. Scenario 3 used MIM experiences the lower loss as the signal drop, the FEC is systematically applied, and the content adaptation is used to reduce the overhead introduced by FEC packets.

Packet losses are translated into video frames dropping due to the cancellation of video-integrated error resiliency. Important packet losses with Scenario 1 are quite predictable since the server is not aware of network conditions and it does not perform any adaptation.

Scenario 2 shows better results than scenario 1 since the server uses the measured packet loss ratio reported by regular RTCP feedback to dynamically adapt the streaming process to network conditions. We can see that there is a significant oscillation in packet loss ratio as a consequence of the adaptive FEC transmission. In fact, as the loss ratio increases, the server responds by transmitting FEC packets to limit the experienced loss rates. However, as the server is unaware of current link layer measurements, it stops sending FEC packets and this leads again to significant packet loss. In contrast, the server in Scenario 3 keeps sending FEC packets since the measured signal quality is poor (bad). This allows the server to anticipate on packet losses that are most likely to occur frequently during poor link-quality periods. Furthermore, with the help of the video adaptation “*transrating*,” the traffic sent by the server is not increased by the FEC redundant packet. This allows maintaining the same achieved throughput at the server side and avoiding the increase in the throughput which in turn may increase the packet losses. Fig. 12 shows the instantaneous throughput of the video traffic achieved at the server side before its transmission over the network. Scenarios 1 and 3 achieve a stable throughput during the experiment, while scenario 2 increases the throughput due to important transmission of FEC packets. The effect of the video adaptation in scenario 3 using

MIM cross-layer adaptation is two folds: First, anticipating on eventual packet loss in the network by continuously monitoring the RSSI measurements. Second, the traffic transmitted by the server is maintained at the same level and aligned with the original video throughput. Thus, the FEC packets added by the adaptation mechanism do not contribute to increase the overall throughput of the video server. This reduces packet losses and enhances the overall video quality.

To assess the user-perceived QoS during link degradations, we use two relevant metrics to measure the objective video quality: Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index (SSIM) [37]. The PSNR estimates the received image quality compared to the original image, while the SSIM measures the structural similarity between the original and the received image. The SSIM method has proved to be more accurate in respect to the Human Visual System (HVS). However, PSNR is widely used for measuring picture quality degradation based on mathematical analysis that proved to be fairly proportional to the human-perceived quality performances. It is derived from the root mean squared error. The PSNR for a degraded $N_1 \times N_2$ 8-bit image f compared to an original image f is computed according to (3) as follows:

$$\text{PSNR} = 20 \times \log_{10} \frac{255}{\sqrt{\frac{1}{N_1 N_2} \sum_{x=0}^{N_1-1} \sum_{y=0}^{N_2-1} [f(x, y) - f'(x, y)]^2}}. \quad (3)$$

Fig. 13(a) and (b) show PSNR and SSIM results for scenario 1 and scenario 3 between frame #1000 and frame #2500. These frames are transmitted within the time interval elapsed between time $t = 40$ seconds and time $t = 100$ seconds which correspond to the poor link quality period where the signal quality is in its poor value (RSSI of 10%). Scenario 1 has a very

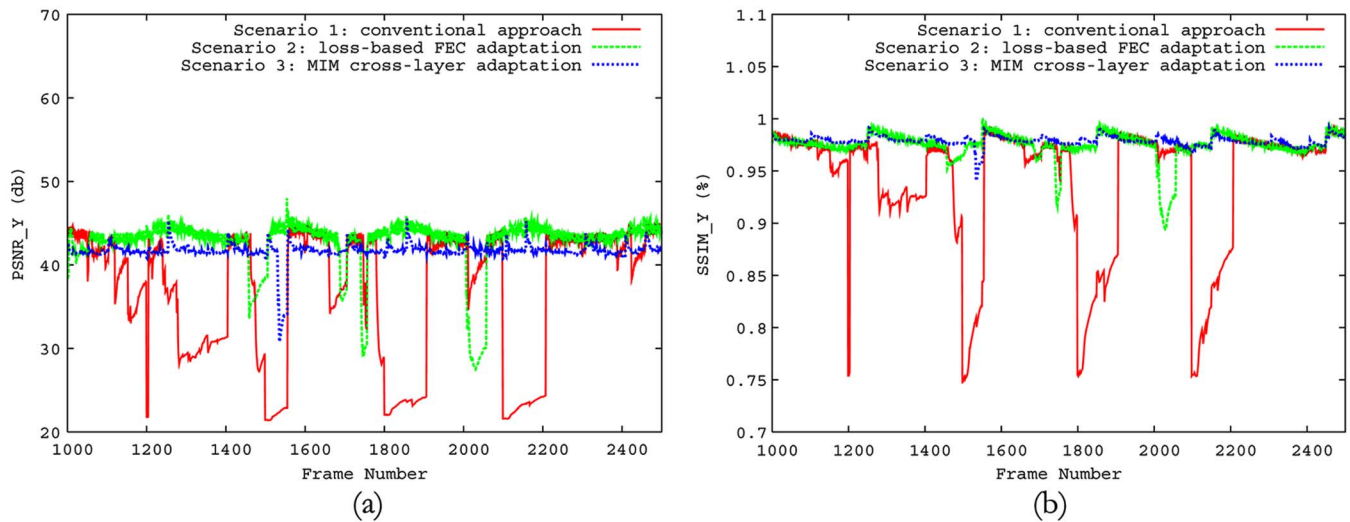


Fig. 13. Perceived video quality (PSNR) measurements for the different scenarios (PSNR and SSIM).



(a)



(b)



(c)



(d)

Fig. 14. Received video quality for Frame 1500. (a) original frame, (b) scenario 1, (c) scenario 2, and (d) scenario 3.

poor PSNR and SSIM quality. This is justified by the excessive packet loss ratio experienced during the aforementioned period. Scenario 2 produces an enhanced PSNR compared to scenario

3, though its value is oscillating during the period by reaching a minimum value of 28 dB. These oscillations affect considerably the subjective video quality.

The video adaptation performed by Scenario 3 affects slightly the PSNR values that keep stable during the critical period and maintain a good subjective video quality. These results are confirmed with the SSIM measurements. In fact, the SSIM of scenario 3 show better performances than scenario 2 and maintain a high level of structural similarity with an index near to 97%. These measurements demonstrate the benefits of MIM cross-layer adaptation in scenario 3 compared to 1 and 2. The received video of these different scenarios can be found in [36]. Coordination between different link quality measurements with different time-scales improves clearly the responsiveness and efficiency of streaming systems over wireless networks.

The snapshots of the received video for the original video [cf. Fig. 14(a)], scenario 1 [cf. Fig. 14(b)], scenario 2 [cf. Fig. 14(c)], and scenario 3 [cf. Fig. 14(d)] are compared for frame #1500. A significant enhancement in the perceived quality is noticeable when MIM cross-layer adaptation is applied.

Although the video is transrated in scenario 3, we cannot notice a difference between the original frame [cf. Fig. 14(a)] and the frame of scenario 3 [cf. Fig. 14(d)] using the original frame size [368×242 pixels]. However, the effect of packet loss on the objective video quality is clear in scenario 1 [cf. Fig. 14(b)] and scenario 2 [cf. Fig. 14(b)].

V. CONCLUSION

This paper introduced a new approach called Meet-in-the-Middle (MIM) for QoS-aware cross-layer design that conciliates both top-down and bottom-up interactions into an adaptive streaming system. MIM consists in gathering service-level characteristics (user, content, and terminal characteristics) and network-level status information into a single rich and dynamic context associated with each active video streaming session. Based on such context information, the server is able to perform a wide range of adaptive QoS mechanisms from content adaptation to PHY/MAC rate adaptation. In addition to the vertical cross-layer interactions that take place at each end-system, our approach involves horizontal communications between end-systems to enrich multimedia session context at the server side using additional measurements carried out at the receiver side. Performance evaluation of MIM cross-layer adaptation revealed a significant improvement of perceived quality compared to conventional approaches. In particular, combining different QoS metrics from different protocol layers proved to be useful in anticipating link quality degradations and thus increasing the responsiveness of the adaptation mechanisms. As future work we intend to develop analytical models for combining network-level metrics into meaningful network condition aggregates which can be used to improve the effectiveness of the adaptive streaming system. Another research direction is to further extend multimedia session context to include other information such as service level agreements. This would allow the service provider to offer a finer granularity of QoS guarantees at different prices.

REFERENCES

- [1] Information Technology—Multimedia Framework (MPEG-21)—Part 7: Digital Item Adaptation 1st ed. Oct. 2004, ISO/IEC Std. 21000-7:2004.
- [2] R. Stewart *et al.*, "RFC 2690: Stream control transmission protocol," *Request For Comments, IETF*, Oct. 2000.
- [3] E. Kohler, M. Handley, and S. Floyd, "RFC: 4340 datagram congestion control protocol (DCCP)," *Request For Comments, IETF*, Mar. 2006.
- [4] L.-A. Larzon *et al.*, "RFC 3828 : The lightweight user datagram protocol (UDP-Lite)," *Request For Comments, IETF*, Jul. 2004.
- [5] H. Schulzrinne *et al.*, "RTP: A transport protocol for real-time applications," *IETF RFC 3550 (Obsolete 1889)*, Jul. 2003.
- [6] D. Clark and D. Tennenhouse, "Architectural considerations for a new generation of protocols," in *Proc. of ACM SIGCOMM'90*, Philadelphia, PA, Sept. 1990, pp. 200–208.
- [7] M. Handley *et al.*, "TCP friendly rate control (TFRC): Protocol specification," *IETF RFC 3448*, Jan. 2003.
- [8] C. Verikoukis *et al.*, "Cross-layer optimization for wireless systems: A European research key challenge," *IEEE Commun. Mag.*, vol. 43, no. 7, pp. 1–3, Jul. 2005.
- [9] S. Shakkottai *et al.*, "Cross-layer design for wireless networks," *IEEE Commun. Mag.*, vol. 41, no. 10, pp. 74–80, Oct. 2003.
- [10] M. Van Der Schaar *et al.*, "Cross-layer wireless multimedia transmission: Challenges, principles, and new paradigms," *IEEE Wireless Commun. Mag.*, vol. 12, no. 4, pp. 50–58, Aug. 2005.
- [11] Q. Zhang, W. Zhu, and Y.-Q. Zhang, "End-to-End QoS for video delivery over wireless internet," *Proc. IEEE*, vol. 93, no. 1, pp. 123–134, Jan. 2005.
- [12] S. Shakkottai, T. Rappaport, and P. Karlsson, "Cross-layer design for wireless networks," *IEEE Commun. Mag.*, vol. 41, no. 10, pp. 74–80, Oct. 2003.
- [13] J. Chen, T. Lv, and H. Zheng, "Cross-layer design for QoS wireless communications," in *Proc. Int. Symp. Circuits and Systems, ISCAS*, 2004, pp. 176–220.
- [14] L. T. Butala, "Cross-layer design for medium access control in CDMA Ad hoc networks," in *EURASIP J. Appl. Signal Process.*, 2005, vol. 2005, no. 2, pp. 129–143.
- [15] S. De, C. Qiao, D. A. Pados, M. Chatterjee, and S. J. Philip, "An integrated cross-layer study of wireless CDMA Sensor Networks," *IEEE J. Select. Areas Commun.*, pp. 1271–1285, Sep. 2004.
- [16] S. Krishnamachari, M. V. D. Schaar, S. Choi, and X. Xu, "Video streaming over wireless LANs: A cross-layer approach," in *Proc. 13th Int. Packetvideo Workshop 2003*, Nantes, France, Apr. 2003.
- [17] Y. Shan and A. Zakhori, "Cross layer techniques for adaptive video streaming over wireless networks," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2002, vol. 1, pp. 277–80.
- [18] T. Ahmed, A. Mehaoua, R. Boutaba, and Y. Iraqi, "Adaptive packet video streaming over IP networks: A cross-layer approach," *IEEE J. Select. Areas Commun.*, vol. 23, no. 2, pp. 385–401, 2005, in the.
- [19] 4G MC-CDMA Multiple-Antenna System on Chip for Radio Enhancements [Online]. Available: <http://www.ist-4more.org> IST-507039
- [20] Jointly Optimizing Multimedia Transmission in IP- Based Wireless Networks [Online]. Available: <http://www.ist-phoenix.org> IST-001812
- [21] Network of Excellence on Wireless Communications [Online]. Available: <http://newcom.ismb.it/public/index.jsp> IST-507325
- [22] [Online]. Available: <http://www.enthrone.org/IST-507637>
- [23] M. Handley and V. Jacobsen, "SDP: Session description protocol," *RFC 2327, IETF*, Apr. 1998.
- [24] T. Guenkova-Luy *et al.*, "Harmonization of session and capability descriptions between SDPng and MPEG-21 digital item adaptation," *IETF Draft, Work in Progress*, Feb. 2005.
- [25] J. T. Haratcherev, K. Langendoen, R. Lagendijk, and H. Sips, "Automatic IEEE 802.11 rate control for streaming applications," *Wireless Commun. Mobile Comput.*, vol. 5, pp. 412–437, 2005.
- [26] T. Braskich, "Optimization of a link adaptation algorithm for voice over wireless LAN applications," in *Proc. IEEE WCNC'05*, New Orleans, LA, Mar. 2005, vol. 3, pp. 1602–1607.
- [27] D. Qiao *et al.*, "Goodput analysis and link adaptation for IEEE 802.11a wireless LANs," *IEEE Trans. Mobile Comput.*, vol. 1, pp. 278–292, Dec. 2002.
- [28] J.-L. C. Wu *et al.*, "An adaptive multirate IEEE 802.11 wireless LAN," in *Proc. 15th Int. Conf. Information Networking*, Oita, Japan, Jan. 2001, pp. 411–418.
- [29] M. Li *et al.*, "Weather forecasting—Predicting performance for streaming video over wireless LANs," in *Proc. 15th ACM NOSSDAV'05*, Stevenson, WA, Jun. 2005.
- [30] I. Djama and T. Ahmed, "A cross-layer interworking of DVB-T and WLAN for mobile IPTV service delivery," *IEEE Trans. Broadcast.*, vol. 53, no. 1, pt. 2, pp. 382–390, Mar. 2007.
- [31] H. Ernst *et al.*, "Transport layer coding for the land mobile satellite channel," in *Proc. IEEE VTC'04*, Milan, Italy, May 2004, vol. 5, pp. 2916–2920.

- [32] A. Nafaa *et al.*, "Joint loss pattern characterization and unequal interleaved FEC protection for robust H.264 video distribution over wireless LAN," *Comput. Networks J.*, vol. 49, no. 6, pp. 766–786, Dec. 19, 2005, et.
- [33] Joint Scalable Video Model (JSVM). Hangzhou, Japan, Oct. 2006, 8.0, ITU-T VCEQ JVT-U202.
- [34] VideoLan Open Source Project [Online]. Available: <http://www.videolan.org/vlc/>
- [35] MadWiFi Open Source Driver [Online]. Available: <http://madwifi.org/>
- [36] Download Page: Received Video Sequences [Online]. Available: http://www.labri.fr/perso/djama/ieee_tm.htm
- [37] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004.



Ismail Djama (M'07) received the computer engineer degree from the Institut National d'Informatique, Algiers, Algeria, in 2003 and the master's degree in computer science from University of Besançon in 2005. He is pursuing the Ph.D. degree in computer science at CNRS-Lab LaBRI, University of Bordeaux, Talence, France.

His main research interests are the quality of service in next generation networks, cross-layer optimization, universal media access and delivery, multimedia adaptation, and multimedia protocols. Currently, he is working on European project ENTHRON and different projects with CNES (Centre National d'études spatiales), Toulouse, France.



Toufik Ahmed (M'04) received the computer engineer degree (Honors) from the Institut National d'Informatique, Algiers, Algeria, in 1999 and the Master's and Ph.D. degrees in computer science from the University of Versailles, Versailles, France, in 2000 and 2003, respectively.

He was a Visiting Scientist at the School of Computer Science, University of Waterloo, Waterloo, ON, Canada, in 2002 and a Research Fellow at PRISM Laboratory, University of Versailles, in 2003–2004. He is currently an Associate Professor at ENSEIRB

School of Engineers/University of Bordeaux I, Bordeaux, France. His main research activities concern quality of service for multimedia wired and wireless networks, cross-layer optimization, and end-to-end signalling protocols. His

work on quality of service and video delivering has led to many publications in major journals and conferences. He has worked on a number of projects including the IST IP ENTHRON project funded by the European Commission.



Abdelhamid Nafaa (M'05) received the Masters and Ph.D. degrees in 2001 and 2005, respectively, from the University of Versailles, Versailles, France, where he was involved in several national and European projects including NMS, IST-ENTHRON1, IST-ATHENA, and IST-IMOSAN.

He is a Marie Curie Research Fellow under the EU-FP6 EIF Marie Curie action that seeks broader synergy in the European research space. He has been granted the Marie Curie award to undertake research work at UCD in the area of multimedia services distribution over carrier-grade networks.

Before joining UCD, he was an Assistant Professor at the University of Versailles and acted as Technology Consultant for U.S. and a European-based companies in the area of reliable multimedia communication over WiFi technology and IMS-based multicasting in DVB-S2 satellite networks, respectively. Recently, he has been involved in writing a successful EI-funded proposal VidAs that aim to develop P2P-based VOD services distribution for network operators. He is also involved in a successful FP7 proposal CARMEN that aim to develop a mixed WiFi/WiMax wireless mesh networks to support carrier-grade services. He is a co-author of over 25 technical journal or international conference papers on multimedia communications.



Raouf Boutaba (SM'01) received the M.Sc. and Ph.D. degrees in computer science from the University Pierre and Marie Curie, Paris, France, in 1990 and 1994, respectively.

He is currently a Professor of computer science at the University of Waterloo, Waterloo, ON, Canada. His research interests include network, resource, and service management in multimedia wired and wireless networks.

Dr. Boutaba is the founder and Editor-in-Chief of the IEEE TRANSACTIONS ON NETWORK AND SERVICE MANAGEMENT and is on the editorial boards of several other journals. He is currently a distinguished lecturer of the IEEE Communications Society, the chairman of the IEEE Technical Committee on Information Infrastructure and the IFIP Working Group 6.6 on Network and Distributed Systems Management. He has received several best paper awards and other recognitions such as the Premier's Research Excellence Award.

Hybrid Overlay Networks Management for Real-Time Multimedia Streaming over P2P Networks

Mubashar Mushtaq and Toufik Ahmed

CNRS LaBRI Lab. – University of Bordeaux 1
351 Cours de la Libération, Talence Cedex 33405 – France
{mushtaq, tad}@labri.fr

Abstract. Recent growth of the multimedia content delivery over the Internet and the popularity of the peer-to-peer (P2P) architecture have opened new horizons for emerging novel services over the Internet. Currently, most of multimedia services are being offered to the end users by using set-top boxes installation on the client's premises, with integrated media storage capabilities and their adaptation. The organization of the end-clients in P2P fashion has great potential to change business models to offer new value-added multimedia services and therefore to generate substantial revenue for service providers. In this paper, we present a mechanism to organize the sender peers in hierarchical hybrid overlay networks. The objective of such organization is to facilitate the receiver peer (content consumer) to select best sender peers for the provision of better QoS (Quality of Service). To construct the hybrid overlay networks, peers offering the same video quality are placed together at the same level of overlay networks. The organization of sender peers within these overlays is subject to (1) the semantic of the video provided by the peer (base layer, or enhancement layers) and (2) the QoS offered by each peer along the end-to-end path. The proposed streaming mechanism is receiver-centric where receiver peer selects a number of sender peers from the overlay networks to receive media contents. The performance evaluation performed using ns-2 simulator shows that hybrid overlays organization mechanism is helpful to enhance the overall QoS by significant improvement in received video packets throughput, the packets drop ratio and transmission delay.

1 Introduction

P2P networking frameworks have recently received a lot of attention by the research community, as they provide an efficient infrastructure to use available networking resources in a more transparent, scalable and cost-effective way. P2P networks were initially designed and were considered suitable for huge contents distribution across networks. These networks possess distributed, scalable, cost-effective, cooperative resource sharing, self-organizing, and many more characteristics that have encouraged service providers (SPs) to deploy many real-time applications over large scale heterogeneous networks. This is considered as a giant shift of technology that has changed the way people interact with the technologies. The most important real-time

applications based on P2P framework include: video streaming, video on demand, P2P based IP-TV, P2P gaming, signaling for IP telephony, etc. Overlay networks are considered as the most promising infrastructure for the deployment of distributed applications. Many existing multimedia applications over P2P networks are based on the organization of the participating peers in overlay networks. Many research developments have been observed overlay networks organization that address the scalability, efficient management, and self organization of overlay networks. The well known overlay networks like Content Addressable Network (CAN) [1], Chord [2], and Pastry [3] have been proposed for the Internet routing to guarantee better quality of service.

On the other hand, recent advances in audiovisual content coding have favored a widespread growing of multimedia streaming and Video on Demand (VoD) services. A number of home holders are subscribing to these services by the service provider through a set-top box. These boxes are generally equipped with specific functions such as storage space, on which the client can store programs, movies, and other desired contents. The service provider has full access for indexing, tracking and to use the contents available in these set-top boxes. These characteristics enable service providers to use such boxes as independent content servers. In the rest of paper, we refer these boxes as peers when used in P2P networks. The terms set-top-box and peer are used interchangeably. In fact, most of these boxes are capable of performing media content adaptation in accordance with the end client's requirements. These features extend the usability of these boxes to that of media servers for the end clients possessing different characteristics, i.e. heterogeneous clients. Another point of heterogeneity is related to the end-user connection. Most of the end-clients (service subscribers) connect to the network using ADSL (Asymmetric Digital Subscriber Line) or CMTS (Cable Modem Termination System), where uplink capacity is 3-8 time lesser than that of the downlink. If media content is requested from a single box, it cannot be achieved. In fact, a single peer might not be able to meet the requirements of any one request. In this regard, we propose to select multiple sender peers to cooperate in the streaming mechanism using multi-source streaming. The use of multiple peers for the service delivery reduces the load on the central server, reduces the start-up delay, latency, and improves the overall Quality of Service (QoS) of the system.

In this paper, we present a quality adaptive mechanism for the multimedia streaming and video on demand services over P2P networks. This mechanism is based on the organization of sender peers in hybrid overlay networks. These overlay networks are formed on the basis of offered video quality and end-to-end probing among the sender and receiver peers.

The network topology for multimedia streaming and video on demand services is illustrated in the Fig. 1 where different service clients with variant characteristics are connected to the network. In this topology, we consider only those peers which have the requested media contents and are willing to participate in the streaming mechanism. The receiver peer selects multiple sender peers to receive the media contents from the overlay networks. The received media contents are combined to decode the received video with a higher quality. The sender peers are selected from the different overlays on the basis of receiver characteristics and its preferences.

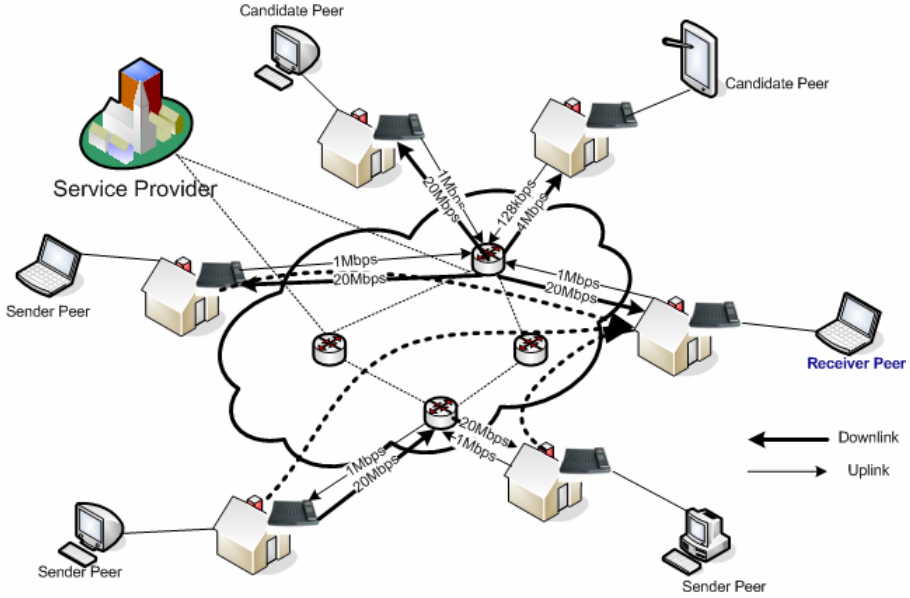


Fig. 1. Network Topology for Multimedia Streaming / Video on Demand

The rest of the paper is organized as follows. A brief motivation and some of the related work is presented in the section 2. Section 3 describes an overview of the scalable video coding (SVC) scheme. The proposed streaming mechanism based on the organization of sender peers in hierarchical overlay networks is given in section 4. The evaluation results of proposed mechanism using simulations are presented in section 5. Finally, concluding remarks are presented in section 6.

2 Motivation and Related Works

There have been significant studies carried out in the context of multimedia streaming applications and for the construction of overlay networks to enhance the QoS. Most of these studies have investigated and proposed solutions for the efficient routing algorithms, organization techniques, and self organization.

Zhang et al. [4] developed a framework for live media streaming that is based on data-driven overlay networks where each node periodically exchanges data availability information with other nodes. The management of nodes which join the networks is controlled by an origin node that is persistent during the life time of the streaming session. In the proposed mechanism, a scheduling algorithm is devised for heterogeneous clients. The nodes in the data-driven overlay network are organized in a Breath-First-Search (BFS) tree, where origin node is located at level '0' while any node at level 'k' can be reached in 'k' hops from the origin.

Tian et al. [5] presented a framework for the Hybrid Overlay Structure that is based on Random Walks. The proposed organization is locality-aware where nodes are organized in different clusters. Crespo et al. [6] proposed the “SON” mechanism for the organization of different nodes. The proposed mechanism is based on the semantic based organization of nodes. Nodes having the same type of media files are organized close to each other and therefore clustered together. The semantic based organization result into improvement of the system performance in terms of the efficiency in query search.

The organization of peers (nodes) in overlay networks systems also has been presented in research work such as CAN [1], CHORD [1], Pastry [2], and Tapestry [7]. These systems provide efficient routing mechanism for the structured P2P overlay networks. The major proposed solutions organized peers in overlay networks on the basis of distances, i.e. Locality-aware. The motivation behind our proposed mechanism is to use the overlay networks for the best sender peers selection in order to support heterogeneous clients and to improve the QoS while efficiently utilizing the available network resources. In our proposed hybrid overlay networks mechanism, we organize different sender peers in different hierarchical overlay networks considering both (1) the semantic of the video and (2) the QoS offered by each peer. The detailed hybrid overlay organization mechanism is presented in section 4.

3 Scalable Video Coding (SVC)

Multimedia content adaptation is considered as an important technique to provide QoS management for multimedia delivery over heterogeneous networks to meet the requirements of clients with distinct characteristics. Many video encoding schemes have been proposed for real-time applications operating for heterogeneous networks and terminals. Layered encoding and multiple description coding [8] are considered suitable for many applications including P2P services. At present, Scalable Video Coding (SVC) is considered most promising video encoding format for streaming application in heterogeneous networks and terminals [9]. A scalable video coding is used to produce highly compressed bit-streams, to generate a wide variety of bit-rates. An original SVC stream can be truncated to produce videos of different qualities, resolutions, and frame rates using respectively SNR (signal-to-noise ratio), spatial, and temporal scalabilities.

In this paper, we focus on the use of SVC video formats as an important component of the proposed P2P system while organizing the sender peers in hybrid overlay networks. The SVC characteristics make it more suitable for heterogeneous environments especially P2P networks. A layered stream representation of SVC in terms of spatial, temporal, and SNR resolution is shown in Fig. 2. In SVC encoding scheme, each quality tier is decodable with different characteristics. The first tier providing the basic quality of the video is called “Base Tier” while other tiers which are used to enhance the overall video quality of the base tier are called “Enhancement Tiers” [10].

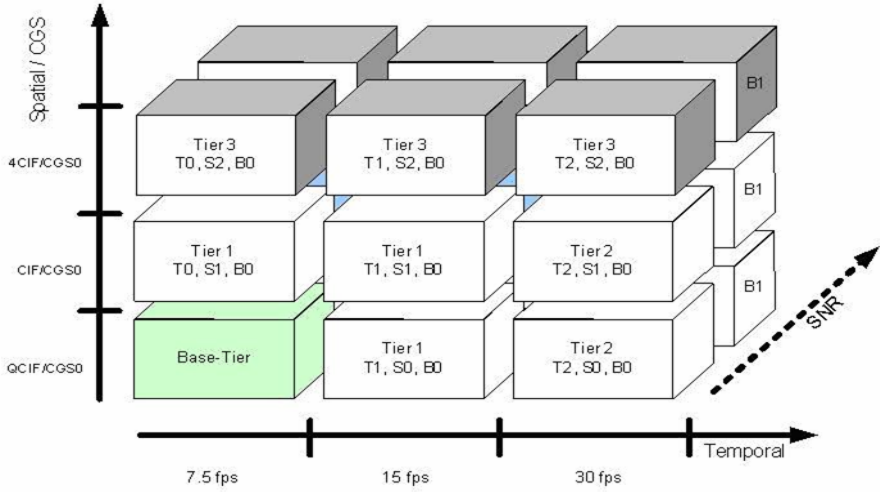


Fig. 2. Scalable Video Encoding – A Layered Stream Representation

4 Hybrid Overlay Networks Organization: Semantic and QoS-Awareness

In this section, we describe our proposed multimedia streaming mechanism over P2P networks. Our targeted P2P streaming solution is based on the collective communication among multiple senders towards a single receiver where the receiver peer orchestrates the overall streaming mechanism. Receiver peer maintains the list of the entire candidate peers which have the requested media contents and are ready to participate in the streaming process. A peer is considered as a candidate peer if it has the requested contents and it is willing to participate in the streaming mechanism. The sender peers are organized in different hybrid overlay networks based on the offered content (i.e. which SVC tier is provided per peer) and based on the QoS offered by the peer. The organization of peers in hybrid overlay networks and their management is described in the following sub-sections.

4.1 Overlay Networks Organization

Many overlay organization schemes have been proposed to address different issues including efficiency, scalability, self-organization, efficient routing. For this study, we have slightly different objective for the organization of the sender peers in overlay networks. We aim to facilitate the receiver peer to select the best sender peer to enhance the overall QoS for the offered services. In this paper, we organize the sender peers in hierarchical hybrid overlay networks. This hybrid peer organization is based on two important characteristics, (1) semantic aware peer organization, (2) quality aware or locality-aware peer organization. The main goal for this hybrid based organization is to bring sender peers closer to the receiver peer and to provide an efficient mechanism for the selection of best peers with the required video quality.

The overlay networks can be classified on the basis of peers' organization criterion. Here, we described some of these classifications along with our proposed Hybrid peers organization.

- **Locality based peer organization:** Locality based peer organization is performed on the basis of geographical/physical distances between sender and receiver peers. The locality of peers is determined by different parameters, for example, Euclidean distances, logical distances estimated using number of intermediate peers (e.g. number of Hops, TTL measurement, etc.), and using end-to-end round trip time (RTT) measurements. Locality-aware or QoS-aware techniques primarily intend to reduce the delays by minimizing the logical distance between the source peers and the receiving peers for each service request. Such locality based peers organization results in the minimizing of start-up delay and latency during the streaming mechanism. Our mechanism is based on RTT measurement as an indicator for the quality of the link between the sender and the destination.
- **Semantic based peer organization:** The structure of media contents is useful for the selection and organization of the appropriate peers in overlay networks. In semantic based organization, peers with the same media contents or offering the same video quality are placed together in overlay networks. Semantic based peers organization influences the delivery of important parts of the audiovisual content to the important peer. Such organization makes it convenient to search media contents rapidly. Semantic based peer organization can be a good choice for scalable video coding and in other coding schemes where media contents are distinguished according to priority. Our mechanism is best adapted to SVC coding or any schemes that can provide priority between sub-streams.
- **Hybrid peer organization:** We propose hybrid peers organization that is influenced by combining locality based (or quality-based) organization and semantic based organization schemes. In this hierarchical hybrid overlay organization, sender peers are arranged in hierarchical overlay networks. The streaming mechanism is receiver-centric. The sender peers offering the same video quality tiers of the same media types are placed at the same level in overlay networks. Sender peers within each level of overlay networks are organized according to locality-aware peer organization. We use round-trip-time "RTT" as an indicator of the offered QoS by each peer. In fact, the best peer is the one offering the best QoS (i.e. the lower RTT). This peer should be selected by the receiver to contribute the base tier of the video. In [11], we have noticed that "RTT" can be used to improve the overall QoS for the streaming applications in heterogeneous networks like P2P. The motivation behind using RTT is its relationship with the offered bandwidth on the end-to-end path. A simplified version of the equation given the bandwidth (data rate) of a particular TCP-like session is given in Eq. 1. The MTU represents the maximum transfer unit and the "Loss" represents the packet loss rate [12].

$$Bandwidth = 1.3 * MTU / (RTT * \sqrt{Loss}) \quad (1)$$

It is commonly agreed that the lower the “RTT” the more the offered data rate (bandwidth) is higher (c.f. Eq. 1). Thus, peers offering lower RTTs are considered as having a higher available bandwidth and that may be shared fairly among the services. This leads us to propose an efficient peer organization mechanism based on active measurement of peers. Any peer is considered as an important peer if it provides the lowest “RTT” and is assigned to contribute the SVC base video tier.

A hybrid organization of the sender peers in different hierarchies is illustrated in Fig. 3. Here, sender peers are divided into three different levels of overlay networks. The organization of these peers is carried out on the basis of the “RTT” probing and on the basis of offered video quality tiers. Overlay 1 consists of the peers offering the base SVC tier quality of video, overlay 2 and overlay 3 depicts the peers offering enhancement SVC tier 1 and enhancement SVC tier 2 respectively. Within each overlay network, peers are organized according to the end-to-end RTT probing carried out between the receiver and each sender peer. We exploit the “MinHeap” Tree structure for the organization of peers. A “MinHeap” is a minimum complete binary tree where the value in each node is less or equal to those in its children. In such tree structures ‘n’ peers can be organized in $\Theta(n \log n)$ time. The organization of peers inside the “MinHeap” tree structure is shown in Fig. 3. The peer offering minimum “RTT” is always present at the root of the overlay tree and provides an efficient mechanism to select the best peer offering the best.

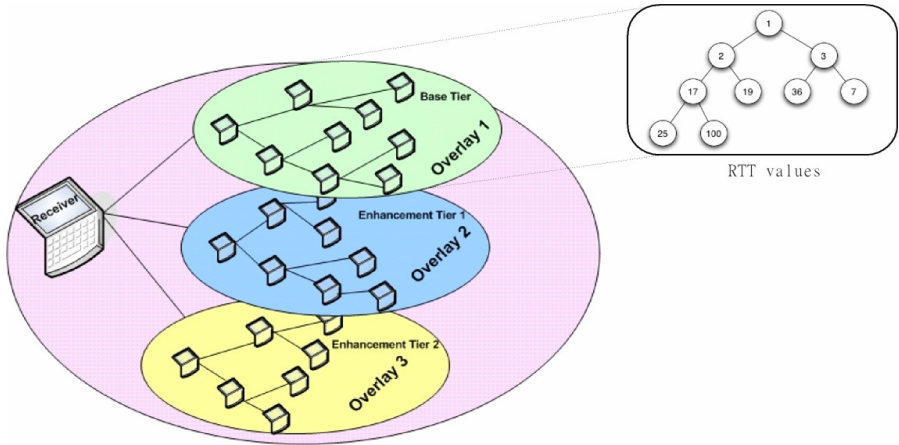


Fig. 3. Layered Organization

4.2 Management of the Sender Peers

The main components of the P2P based media streaming and video on demand services include: indexing of the media contents available to each sender peer,

signaling among the peers for their activation, organization of sender peers in virtual overlay networks, media contents retrieval, selection of best sender peer to serve for subsequent request, sender peers switching, selection of quality adaptation path, etc..

In this paper, we did not discussed the issues of indexing, signaling, and data retrieval mechanisms and we are more concerned with the organization of peers in overlay networks for the efficient sender peers selection to enhance the overall QoS.

The management of senders peers in the P2P streaming mechanism is essential as (1) a single peer is not able to serve alone a service request due to its uplink capacity limitations, (2) some selected peers might leave the P2P network at any moment – for example, when the end-user may switch off its system, and (3) additional contributions from new source peers improves the overall Quality of Service (QoS).

- **Selection and Switching of Sender Peers:** We propose to select multiple sender peers from different overlay networks to address the aforementioned issues. We described earlier that scalable video coding is used for video encoding schemes, thus the selection of the best peers is performed by assigning the best quality tier (Base tier) to the best sender peer. A peer is consider as the best peer if it offers the lowest end-to-end “RTT” to the destination. The peers are organized using the “MinHeap” tree structure in different overlay networks thus the best peer in each overlay is located at the root of the “MinHeap” tree that is accessible to the receiver peer in $\Theta 1$ time.

P2P networks possess highly dynamic nature and show unpredictable behavior; any peer node can enter or leave the network without any prior notification. In this situation, P2P architecture is not considered reliable for multimedia streaming applications which require permanent connection establishment throughout the streaming process. During long streaming sessions, network conditions can change drastically. This may involve many factors among which (1) a sender peer crashes and /or stops contributing to the media content, (2) a shared bandwidth is changed since the network conditions change, (3) some new peers enter the network and provide better bandwidth sharing and low “RTT” value, (4) heavy traffic on a particular peer can cause more packet loss, and (5) high inter packets delay ultimately degrades the overall QoS.

Network fluctuations degrade received video quality dramatically. Such network fluctuations can be detected using the “RTT” measurements. If “RTT” values become high for a certain sender peer, the sender peer assumes that network links are no longer suitable for the streaming session. This issue is handled by the switching mechanism in two ways. First, if there are some peers available with low “RTT” values in the subset of tracked peers, the effected peer could be replaced with this peer to maintain a smooth video packet delivery. In other case, if no other best peer is available, stream switching is enforced instead of peers switching. In stream switching, receiver peer re-adjusts the received video streams dynamically. This can be done easily using SVC stream organization as it offers three-dimensional scalability (i.e. special, temporal and SNR). Thus, the received video quality is adapted to the actual available resources in the network by dropping the streaming of some SVC tiers. We implemented the exponential weighted moving average “EWMA” to prevent from oscillation effects while updating peers after

each “RTT” calculation. EWMA is an Exponential Smoothing technique [13] that employs one exponential smoothing parameter to give more weight to recent observations and less weight to older observations and vice-versa.

Can the best peer always remain the best? Our targeted architecture for the proposed solution is based on many to one streaming, i.e. a single receiver selects multiple sender peers to receive different parts of the requested media contents. In this scenario, what happens when a number of receivers request for the same media contents? Every receiver peer should prefer to select the best sender peer to receive the contents with the highest quality, in this case due to a certain overload on the best peer; can it still remain the best peer? Nash presented in his classical “Nash Equilibrium” theory [14] that no player can end with higher scores if every player of the certain game choose the best policy. We can state that if every receiver peer intends to select the same best peer, no receiver can get the entire video with good quality and there might be a system crash due to the heavy load (in the case of large number of receiver peers). The best peer for a particular session will not then remain the best for the following session.

The proposed peer organization and their selection provide an efficient way of handling this issue. Firstly, our proposed solution is receiver-centric, and secondly it is based on the probing of sender peers using “RTT” as an indicator of the QoS. We noticed that the probed end-to-end “RTT” is not the same for all the receiver peers, and being the receiver-centric solution every receiver maintains the overlays of sender peers and candidate peers in its own prospective. In our mechanism, the best peer for a receiver might be the worst sender peer for other receiver peer.

- **Quality adaptation:** Quality adaptation determines the way to assign the video quality tiers to respective sender peers by selecting from different overlay networks. The hybrid overlay organization facilitates to determine the quality adaptation in an efficient manner as all the peers are classified on the basis of offered video quality and QoS. Fig. 4 describes scenarios how a receiver peer can determine quality adaptation based on its preferences. We have presented different quality adaptations patterns (i.e. horizontal adaptation, vertical adaptation, diagonal adaptation, ZigZag adaptation) which represent the patterns to select the sender peers to receive the video quality tiers.

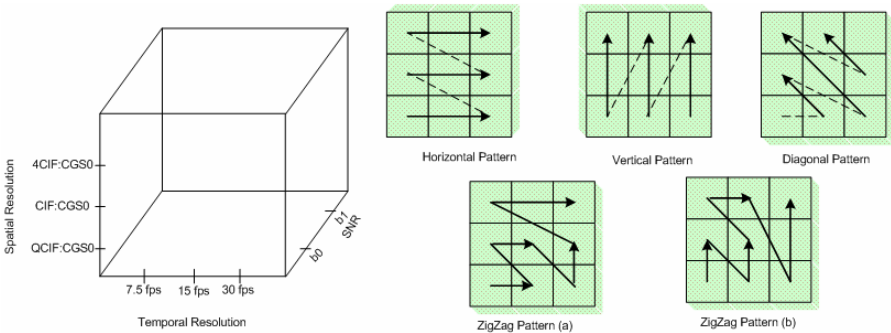


Fig. 4. Quality Adaptation Patterns

Horizontal and Vertical ordering is more suitable for cases where the receiver intends to receive video quality based on a single parameter, i.e. temporal, spatial, or SNR. ZigZag and Diagonal patterns provide the mechanism to select different sender peers to receive video quality based on multiple parameters. We suggest selecting multiple sender peers from the overlay networks based on ZigZag adaptation for our proposed streaming mechanism. In this way, a better QoS for the received video can be ensured on the basis of available resources.

5 Performance Evaluation

This section describes the simulation results for the proposed streaming mechanism performed using the ns-2 simulator.

5.1 Network Model

We consider Gnutella like P2P topology for the simulations and we organize different groups of peers on the basis of proposed hybrid overlay scheme. We distribute the SVC tiers equally among different overlays to examine the performance of the proposed mechanism. A receiver peer intends to receive the real-time video packets from multiple sender peers using P2P network. The receiver peers use “RTT” for the probing of sender peers and best peer selection is performed on the basis of lower “RTT”. “RTT” values are only indicative that reflect the current network conditions and give sufficient information to choose the best peer.

We perform simulations to receive the video quality of CIF/CGS0 with 15 fps. The receiver peer activates a particular sender peer from each overlay group depending on the “RTT” value and the offered video quality based on ZigZag pattern (a) shown in Fig. 4. Each sending peer contributes different quality tiers of the original video file, so that it can be used to reconstruct a video file with the best quality at the receiver node. For the test cases, we generated 4 different quality tiers using MPEG-4 trace files where Base Tier offers 40 % throughput of original video, enhanced tier 1 offers 30 % throughput of original file, enhanced tier 2 offers 20 % throughput of original file and enhancement tier 3 offers 10% throughput of the original video file.

We noticed that no source peer is providing 100% throughput of the original video but if a receiver peer receives all the 4 tiers from different sender peers, it is possible to reconstruct the original media file with 100% quality with the selected characteristics of scalable video coding scheme. The simulation time was 60 seconds and the presented results are the average results of the multiple runs of these simulations. We attached two “CBR sources” to overcharge the network. “CBR source 1” is started at time 5 second and stopped at time 55 second. “CBR source 2” is started at time 10 second, and stopped at time 50. Both sources injected constant throughput of 512 Kbps with 512 Bytes UDP packet lengths. We simulate the adaptation mechanism for two scenarios.

- **Scenario without Quality Adaptation:** Simulation without applying any quality adaptation mechanism. In this case, P2P system works as in downloading modes. The sender peers are selected on a random basis.

- **Scenario with the proposed Quality Adaptation:** Simulation with quality adaptation mechanism is performed by best peer selection and their switching based on the proposed hybrid overlay networks as described in section 4.

5.2 Simulation Results

Fig. 5 shows the received video throughput at receiver peer for both scenarios. We observed that the quality adaptation mechanism improved the received throughput compared to the scenario without quality adaptation. We have noticed few packets drop even with quality adaptation mechanism. These packets drops are caused due to the heavy stress on the network created by CBR/UDP traffic which is presented in Fig. 6. Packet drop ratio is much lesser in the scenario with quality adaptation compared to scenario without adaptation so, our quality adaptation mechanism works fine for this case too. This enhanced throughput and lower packets drops results into the overall improvement in QoS.

We monitored all the active network links constantly and select only one sender peer from each overlay group, i.e. peer having lowest “RTT” and present at the root of “MinHeap” tree. The stream switching is done by selecting the best sender peer offering the better QoS (high bandwidth share). We performed peer switching and implemented EWMA mechanisms to avoid fluctuations caused by peers arrival or removal. Fig. 7(a), Fig. 7(b), Fig. 7(c), and Fig. 7(d) describe the comparison between received and original base tiers, enhancement tier 1, enhanced tier 2, and enhanced tier 3 respectively. We can see clearly that our proposed mechanism performs a smooth video delivery with higher quality, lower loss and delay. A summary of these results is given in Table 1.

The overall received 1-way packets delay (from sender peer to receiver peer) for both cases is presented in Fig. 8. The packets transmission delay has been significantly improved when quality adaptation mechanism is applied. We reconstructed the expected video file using the “AKIYO” video sequence for the received video quality at the receiver end to evaluate the PSNR (Peak signal-to-noise ratio) objective measurements. The comparison results for PSNR for the videos generated in both scenarios to the original video are shown in Fig. 9.

Table 1. A Summary of Simulation Analysis

	Received Video Quality (%)	Average Received Base Tier (%)	Average Received Enh. Tier 1 (%)	Average Received Enh. Tier 3 (%)	Average Received Enh. Tier 2 (%)	Average Packet Delay (ms)
With quality Adaptation	98	99	99	93	99	45
Without Quality Adaptation	76	60	96	57	100	60

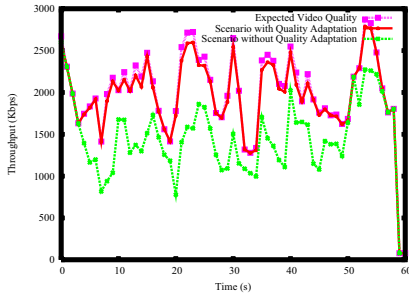


Fig. 5. Received Video Throughput

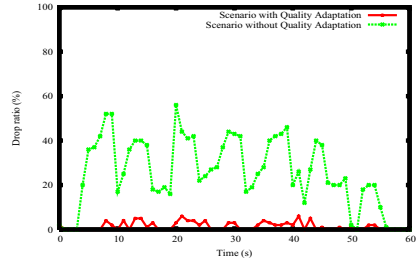


Fig. 6. Packets Drop Ratio

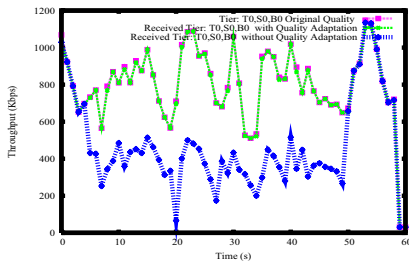


Fig. 7. (a) Base Tier

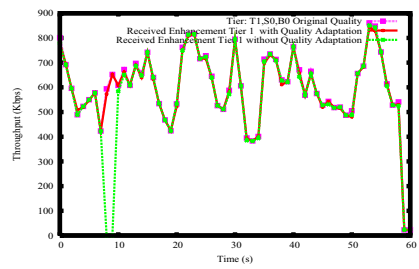


Fig. 7. (b) Enhancement Tier 1

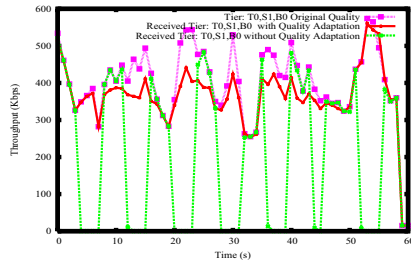


Fig. 7. (c) Enhancement Tier 2

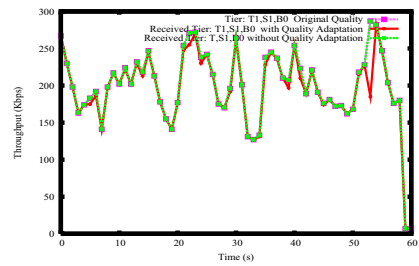


Fig. 7. (d) Enhancement Tier 3

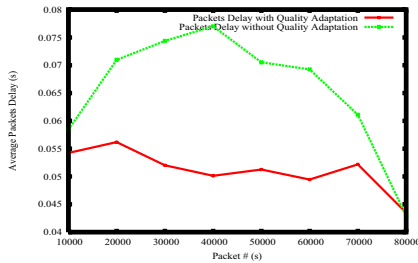


Fig. 8. 1-way Packets Delay

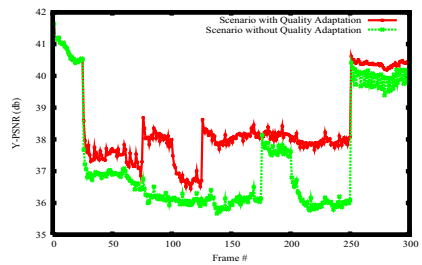


Fig. 9. PSNR

6 Conclusion

In this paper, we presented a hybrid overlay network mechanism for the multimedia streaming and video on demand services in P2P environment. The hybrid overlay networks are constructed on the basis of video quality offered by sender peers and active end-to-end probing of sender peers using “RTT”. The overlay organization provides a mechanism to select the best sender peers in $\Theta 1$ time. The proposed mechanism is evaluated using simulations and a significant improvement in the received throughput especially for the important video quality tiers, lower packets drop ratio for the important quality tiers, and a considerable improvement in the received packets delay are observed.

References

1. Ratnasamy, S., Francis, P., Handley, M., Karp, R., Shenker, S.: A Scalable Content-Addressable Network. In: Proc. ACM SIGCOMM, pp. 161–172. ACM Press, New York (2001)
2. Stoica, I., Morris, R., Karger, D., Kaashoek, M.F., Balakrishnan, H.: Chord: A scalable peer-to-peer lookup service for internet applications. In: Proc. ACM SIGCOMM, ACM Press, New York (2001)
3. Rowstron, A., Druschel, P.: Pastry: Scalable, distributed object location and routing for large-scale peer-to-peer systems. In: proc. of the 18th IFIP/ACM International Conference on Distributed Systems Platforms (Middleware 2001), pp. 329–350 (November 2001)
4. Zhang, X., Liu, J., Li, B., Yum, T.-S.P.: CoolStreaming/DONet: A Data-Driven Overlay Network for Efficient Live Media Streaming. In: proc. of IEEE Infocom 2005, Miami, FL, USA (March 2005)
5. Tian, R., Xiong, Y., Zhang, Q., Li, B., Zhao, B.Y., Li, X.: Hybrid Overlay Structure Based on Random Walks. In: Castro, M., van Renesse, R. (eds.) IPTPS 2005. LNCS, vol. 3640, pp. 152–162. Springer, Heidelberg (2005)
6. Crespo, A., Garcia-Molina, H.: Semantic Overlay Networks for P2P Systems, technical report (last viewed May 15, 2007), available from <http://infolab.stanford.edu/crespo/publications/op2p.pdf>
7. Zhao, B., Kubiawicz, J., Joseph, A.: Tapestry: An infrastructure for fault-tolerant widearea location and routing. Technical report UCB/CSD-01-1141, U. C. Berkeley (April 2001)
8. Vitali, A., Fumagalli, M., Cefriel: Standard-compatible Multiple-Description Coding (MDC) and Layered Coding (LC) of Audio/Video Streams, Internet Draft, Network Working Group (July 2005)
9. Schwarz, H., Marpe, D., Wiegand, T.: SNR-Scalable Extension of H.264/AVC. In: proceedings of ICIP 2004, Singapore (2004)
10. Wu, D., Hou, T., Zhang, Y.-Q.: Scalable Video Coding and Transport over Broadband Wireless Networks. In: Proceedings of the IEEE (September 2000)
11. Mushtaq, M., Ahmed, T.: Adaptive Packet Video Streaming over P2P Networks Using Active Measurements. In: ISCC 2006. proceedings of the 11th IEEE Symposium on Computers and Communications, pp. 423–428. IEEE Computer Society, Los Alamitos (2006)
12. Mahdavi, J., Floyd, S.: TCP-Friendly Unicast Rate-Based Flow Control. Technical note sent to the end2end-interest mailing list (January 8, 1997)
13. Stuart Hunter, J.: The Exponentially Weighted Moving Average. J. Quality Technology 18(4), 203–207 (1986)
14. Nash, J.: Non-Cooperative Games. The Annals of Mathematics 54(2), 286–295 (1951)