

Chapitre x

Distribution Adaptative de Contenus Audiovisuels sur les Réseaux IP

Toufik Ahmed

Laboratoire LaBRI / Université de Bordeaux-1
351 Cours de la Libération,
33405 Talence Cedex,
France.

Téléphone : +33 5 40 00 35 01

Fax : +33 5 40 00 38 95

E-mail : tad@labri.fr

1.1. Introduction

De nouveaux services de distribution de contenus audiovisuels et multimédia fond leur nid sur les réseaux, à la fois fixe (ADSL, câble) et mobile (UMTS, WiFi, DVB-H/T). Cependant, plusieurs caractéristiques de ces réseaux font que le déploiement à grande échelle de tels services présente actuellement un réel challenge par rapport à des applications telles que l'email et le Web. En effet, le protocole IP est en passe de devenir le mode de transport de référence sur ces différents réseaux, ce qui pose certains problèmes liés à la qualité de service (QoS) hérités de ce protocole. Les applications audiovisuelles doivent faire face aux différentes variations de la bande passante et du délai de transfert, tout en restant robuste aux pertes de paquets et aux erreurs de transmission.

Plusieurs avancées technologiques dans le domaine de la compression audio/vidéo (MPEG-4, H.264) associé à la numérisation des infrastructures de distribution de la télévision (DVB) favorisent la convergence des services audiovisuels et multimédia vers IP. Dans ce contexte, un nombre important de réseaux de communications, tels que le WLAN (802.11a/g, HIPERLAN), le WMAN (802.16) et les réseaux cellulaires 2G/3G (GSM, GPRS, UMTS), peuvent être considérés comme candidats pour la distribution des services audiovisuels et multimédia et, 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.

Ces nouvelles technologies constituent également une opportunité pour créer et pénétrer d'autres parts de marché numérique et audiovisuel. Les utilisateurs du service pourront avoir accès à des services nouveaux, des chaînes de télévision aux contenus innovants, des services TV de type pay-per-view interactifs et au format adapté à la taille des écrans utilisés et aux préférences des utilisateurs.

Ces nouveaux services d'agrégation posent la problématique de la QoS. En effet la QoS 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 trafic pour satisfaire l'utilisateur final. Afin de répondre au besoin de l'utilisateur en terme de QoS, 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 usagers et des mécanismes de QoS offerts dans le réseau sous-jacent (profil réseau).

L'interopérabilité de ces éléments sera assurée par un modèle de données en cours de définition dans la norme MPEG-21 et qui rend l'application plus adaptable à son contexte environnant. Cette norme sera employée pour fournir l'appui commun pour mettre en application les fonctionnalités exigées par les applications multimédia (QoS, Sécurité, etc.) et pour assurer le contrôle des ressources dans la chaîne de distribution des services audiovisuels.

Le but de ce chapitre est d'étudier les problèmes liés à la distribution des flux audiovisuels à grande échelle et de présenter une architecture de distribution adaptatif de contenu audiovisuel. Cette architecture adresse plusieurs points liés à la gestion et au contrôle de QoS pour les services audiovisuels transportés sur IP.

Le reste de ce chapitre est organisé comme suit : Section 2 présente les challenges du transport des paquets audiovisuels sur IP. Plusieurs critères liés à la bande passante, taux de perte et le délai ont été étudiés. Section 3 présente le rôle de

la norme MPEG-21 dans la gestion unifiée des services audiovisuels. Section 3 présente une architecture intégrée de distribution adaptative de contenus audiovisuels sur les réseaux IP. Section 4 dresse un panorama des avancées technologiques futures dans le domaine de distribution de contenu audiovisuel. Finalement une conclusion est présentée dans la section 5.

1.2. Exigences et besoins des services audiovisuels

La distribution des services audiovisuels et multimédia sur IP est soumise 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 cette section, nous exposons les principaux problèmes et les solutions, liés à la distribution des paquets audio/ vidéo sur IP trouvés dans la littérature et qui sont résumés dans la Table 1.

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

Table 1: Résumé des problèmes et des solutions de la distribution audiovisuels sur IP.

1.2.1. Gestion de la bande passante

La gestion de la bande passante est liée principalement au choix du codec audiovisuels à utiliser dans le service donné. Ce choix est généralement un compromis entre le taille des unités de données (échantillons), le délai de compression, le débit à la sortie du codec et la puissance du traitement (complexité de l'algorithme). Cette équation à plusieurs inconnus constitue la partie la plus importante dans la conquête du temps réel.

Actuellement, il existe plusieurs normes de codage audiovisuel (voir Figure 1). La dernière en date est la norme H.264 développée par le groupe JVT (Joint Video Team) qui regroupait des experts de l'ISO/IEC et l'ITU-T. Ce codec a été intégré comme partie 10 dans la norme MPEG-4. La norme H.264 vise des services très variés (HDTV¹, VoD², vidéoconférence, etc.). Les tests subjectifs montrent qu'avec

¹ Télévision Numérique Haute Définition (High Definition TV)

deux fois moins de débit par rapport à la norme MPEG-2, H.264 peut atteindre les mêmes qualités visuelles. Ce qui explique son adoption rapide dans différents organismes notamment au sein du 3GPP et DVB pour offrir des services audiovisuels pour les réseaux cellulaires de mobiles et les réseaux de diffusion audiovisuels respectivement.

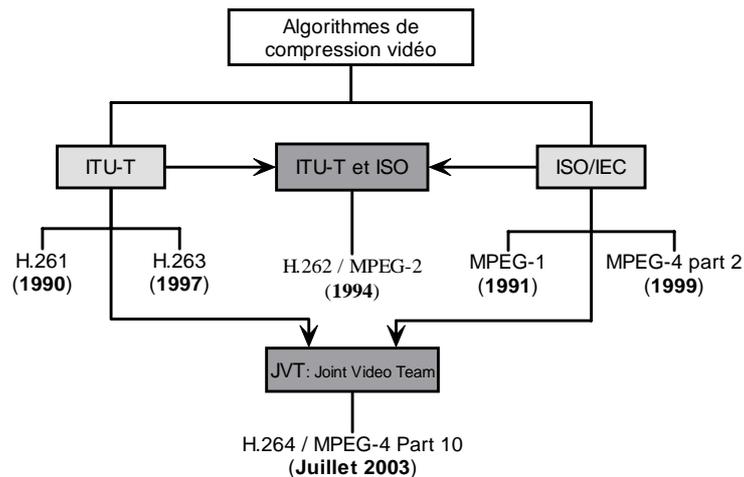


Figure 1: Principales normes de compression audiovisuelles.

Choisir un bon codec répondant au besoin de la bande passante ne suffit pas pour déployer un service audiovisuel sur IP. Il est clair que dans le contexte d'un réseau multiservice, les techniques de garantie de QoS par des mécanismes de planification et dimensionnements traditionnels sont mal adaptées. Vu la nature hautement sporadique et hétérogène des flux transportés (voix, donnée, vidéo, etc.), et les variations rapides des quantités de trafic à transporter, il est nécessaire de mettre en place de mécanismes de contrôle plus dynamiques aussi bien au niveau du réseau qu'au niveau des terminaux. Cette problématique est engendrée par les difficultés de prédire correctement le trafic et donc le comportement du réseau. En conséquence, il n'est pas possible de contrôler le réseau de manière statique en boucle ouverte, mais plutôt requièrent d'observer l'état du réseau pour décider des stratégies de reconfigurations nécessaires ou d'adaptation à effectuer au niveau du trafic audiovisuel.

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 à sa consommation, à travers plusieurs réseaux. D'après ce

² Vidéo à la Demande (Video on Demand)

concept, le niveau de QoS est mesuré dans multiples points: au niveau terminal final (l'utilisateur), à l'entrée et à la sortie de chaque réseau et à la génération du service. Cela permet de déterminer quel point possible dans le réseau est responsable de la dégradation de la QoS et éventuellement proposer des actions pour remédier et / ou compenser cette dégradation.

Ce phénomène s'aggrave avec l'utilisation des services multimédia avec la technique de codage 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). Ce dernier, n'est pas adapté aux services interactifs et conversationnels vu la stratégie employée et qui est basée sur l'utilisation des mémoires tampons pour régulariser et absorber les variations de débit à la sortie du codec. En effet, l'utilisation des 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.

En conséquence, il est nécessaire de gérer la bande passante de manière 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 et au moment de sa connexion il négocie avec celui-ci les paramètres de qualité de service en terme de débit. 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 en dotant la source de mécanismes permettant de détecter les changements des conditions réseau et 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). Une évaluation de cette technique est présentée dans [1] [2]. Cette technique est aussi employée dans les produits RealVideo développés par real network et qui est appelée SureStream [3].
- **SimulCast** (Simultaneous multicast) : dans cette technique la source diffuse plusieurs session 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 à l'autre pour s'adapter à ses besoins en terme 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 passerelle appelée « passerelle multimédia » [5][6]. Le but de cette technique est de changer le débit d'un flux audiovisuels 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, la résolution spatiale et la résolution temporelle [7]. Cette technique est très gourmande en terme de puissance de calcul et de délai de traitement. Elle n'est pas adaptée pour les services directs et interactifs (live) vu les délais supplémentaires imposés par les algorithmes de transcodage.
- **Codage conjoint source/canal** : cette technique peut être utilisée pour les services directs. Elle permet de changer la quantité d'information sortant du codec en changeant le facteur de quantification de la matrice DCT (Discrete Cosinus Transform). Dans cette technique, un module, au niveau de la source audiovisuelle, estime en permanence la qualité du canal de transport (bande passante disponible) et ordonne le module de quantification à augmenter ou à diminuer la valeur du pas de quantification (Quantizer step). 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 résulte en un débit faible à la sortie du codec et une qualité visuelle médiocre, par contre diminuer le pas de quantification résulte en un débit élevé et une qualité visuelle optimale.
- **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 des paquets vidéo. Les envois de paquets vidéo en rafale sont très courants 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 elles-mêmes 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 démunie l'interactivité du service.
- **Streaming audiovisuels avec gestion de bande passante de type TCP-Friendly** : Les applications vidéo utilisent généralement la pile de protocoles 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 à la présence du 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 remédier à ce problème. Il peut être présent à

plusieurs niveaux : 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.). 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.) [8]. Ainsi, il est nécessaire qu'un trafic UDP applique lui aussi un contrôle de congestion pour être compatible TCP (TCP-Friendly) [9]. L'idée de TCP-Friendly est d'empêcher l'application multimédia à rentrer dans une phase où beaucoup de trafic envoyé et peu correctement reçu. Pour éviter ce phénomène, le trafic qui n'est pas compatible TCP peut être supprimé par les routeurs [10], ce qui oblige les applications à opérer en mode contrôle de congestion. La scalabilité audiovisuels 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 générique tels que : LDP « Loss-Delay Based Adaptation Algorithm » [11], RAP (Rate Adaptation Protocol) [12], TFRC « TCP-friendly Rate Control Protocol » [13] qui peuvent être appliqués à cet effet. Le protocole TFRC semble mieux mature et plus robuste. Il utilise un schéma de calcul de débit basé sur l'équation Eq.1. R_{TCP} est le débit autorisé de la source à ne pas dépasser, RTT (Round-Trip Time) est le délai aller/retour, s est la taille du paquet, p est le taux de perte, t_{RTO} est le temps d'attente avant une transmission, b est le nombre de paquet acquitté avec un seul paquet d'acquiescement.

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

La source estime alors le débit disponible dans la liaison et régule son débit par des techniques d'adaptation multimédia (suppression d'une couche, suppression des images B ou P, etc.). Ce type d'estimation de débit et d'adaptation n'est possible que pour les services unicast.

La table suivante (Table 2) résume l'ensemble des techniques présentées précédemment et donne une vue globale sur leur domaine d'application et leur degré de complexité.

Technique	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

Table 2: Complexité et domaine d'application des techniques de gestion de bande passante

1.2.2 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 des pertes de paquets. Un paquet en entier peut être jeté 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 entêtes du paquet, elle peuvent être détectées par les champs de contrôles CRC et le paquet sera rejeté.

Pour combattre les effets des pertes sur le flux audiovisuels décodés, il est nécessaire dans un premier temps de détecter et d'isoler la perte et dans un second temps d'appliquer des techniques de remèdes. La Figure 2 donne un aperçu et une classification des techniques de transport fiable de paquets vidéo.

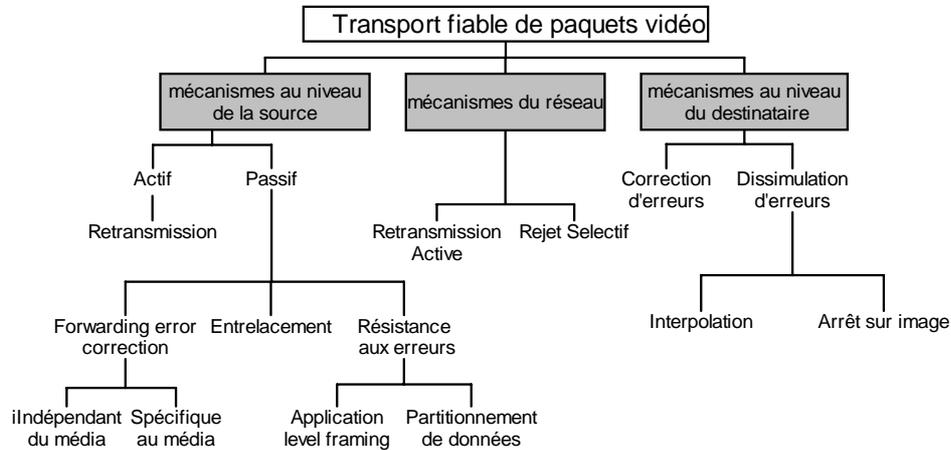


Figure 2: Techniques de contrôle d'erreurs et de pertes

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

- **Retransmission** : dans cette approche, les paquets perdus sont demandés par le récepteur. Un canal de retour est nécessaire pour informer la source quels paquets à envoyer. L'inconvénient majeur de cette technique est l'augmentation du délai de consommation du flux 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 possible en analysant l'historique des délais aller-retour dans le réseau. En utilisant un seuil de tolérance, ce délai est adaptatif en fonction de la charge du réseau. Plusieurs travaux de recherche exploitent cette technique [14][15].
- **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 données de redondances sont dérivées des données originales en utilisant des techniques mathématiques basées sur la théorie du codage [16]. 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 la consommation de bande passante à cause de l'augmentation de l'overhead

des données de redondances. Un deuxième inconvénient est que cette technique impose un retard supplémentaire à cause du stockage des données avant leur traitement. 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 [17][18][19]. Il est ainsi possible de mieux protéger les images I par rapport aux images P, eux même par rapport aux images B.

- **Entrelacement des données** : cette technique vise à réduire l'effet des pertes de paquets en rafale. Si une perte a eu lieu sur un paquet donné, il y a une forte probabilité que le paquet suivant sera lui aussi perdu. Il est mieux de disperser ces pertes sur la totalité du flux audiovisuels par rapport à une perte localisée sur un ensemble d'échantillon donné de sorte que des paquets adjacents soient séparés par une distance garantie. L'inconvénient de cette technique 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.
- **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 de paquet. Pour minimiser la propagation des erreurs vers d'autres flux, il est nécessaire de concevoir un format spécifique de charge utile (payload) pour chaque norme de codage audiovisuel. La source peut organiser mieux 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 une taille de MTU et indépendamment décodable. La plupart des codecs audiovisuels tel 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 techniques de dissimulation des erreurs en estimant l'information perdue. La technique la plus utilisée est de cacher le bloc perdu par un bloc équivalent noir ou vert. D'autres techniques plus adaptées au 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 sont mieux appliquées sur des

pertes isolées mais le problème peut s'avérer plus contraignant si plusieurs paquets successifs sont perdus.

La table suivante (Table 2) résume l'ensemble des techniques présentées précédemment et donne une vue globale sur leur domaine d'application et leur degré de complexité.

Technique	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

Table 3: Complexité et domaine d'application des techniques de gestion des pertes et des erreurs

1.2.3. Gestion du délai et de la gigue

Le délai est le temps nécessaire pour le transfert de bout en bout d'un paquet audiovisuel depuis sa génération à sa consommation par le client. La majeure partie du délai est due au délai du transfert dans le réseau, mais il inclut d'autre délai à prendre en considération tels quel : les délais de numérisation, de compression, et de mise en paquets. La gestion de ce délai est principalement assurée par les mécanismes de qualité de service réseau. 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 du marché dans le déploiement immédiat d'une solution de QoS dans le réseau Internet. Contrairement à la gestion flux par flux de RSVP, les réseaux Diffserv classent les paquets dans l'une des classes agrégées identifiées par le champs DSCP (DiffServ CodePoint) dans l'entête du paquet IP. Dans chaque routeur Diffserv, un comportement particulier est donnée 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 et la somme des comportements individuels de chaque

routeur appartenant à des domaines administratifs. La problématique principale dans l'architecture Diffserv est la gestion et la garantie d'un comportement donné de bout en bout. Le paquet IP peut traverser plusieurs domaines administratifs et le contrat de service ne peut 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 résoudre cette problématique et principalement définir un format de contrat et un protocole de signalisation générique pour négocier et échanger les contrats de services.

La conséquence directe du délai 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 est 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 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 challenge réel pour le déploiement des services audiovisuels. Utiliser une mémoire tampon trop grande rallonge le délai de transfert dans le réseau et rend le service non interactif et non conversationnel. De l'autre côté, utiliser une mémoire tampon trop petite ne permet pas d'absorber la variation de délai de gigue et augmente les taux de pertes des paquets. La Figure 3 illustre le compromis à prendre en considération pour dimensionner la mémoire tampon. Pour un taux de perte de l'ordre de 3%, il faut utiliser une mémoire tampon de 100ms pour pouvoir accepter les paquets arrivant en retard. Ceci rallonge le délai total du service de 180ms. Plusieurs travaux de recherches existent pour définir des modèles adaptatifs et analytiques pour déterminer la taille optimale du buffer de gigue, pour plus de détails voir [20][21][22] et [23].

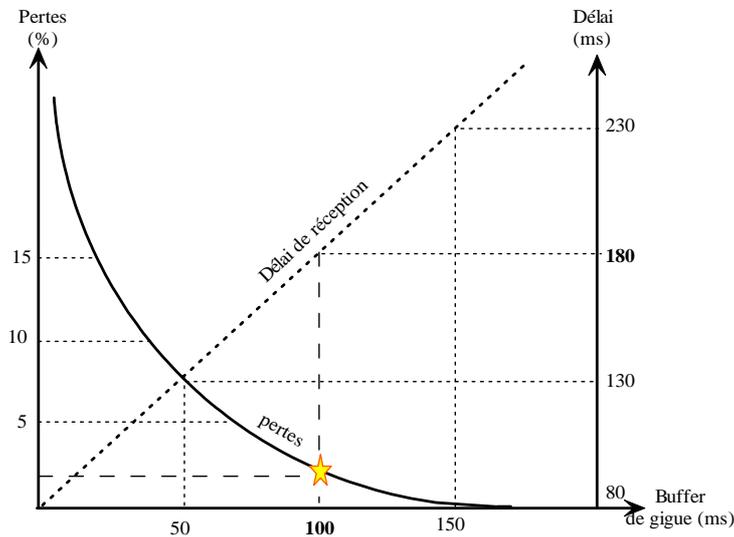


Figure 3: Compromis délai / pertes pour un service multimédia interactif

1.3. Gestion des services audiovisuels : rôle de MPEG-21

Actuellement, beaucoup d'éléments existent pour établir une infrastructure de distribution et de consommation des contenus multimédias. Cependant, il n'existe aucune vue d'ensemble pour décrire comment ces éléments, existant ou en cours de développement, se relient et s'intègrent entre eux pour assurer la meilleure qualité de service à l'utilisateur final. C'est cette problématique qui 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. Son résultat est un cadre ouvert pour la distribution et la consommation de contenu multimédia sur l'ensemble de la chaîne du service, depuis la génération du contenu à sa consommation par les clients.

Sept éléments ont été identifiés dans l'architecture MPEG-21, voir Figure 4. 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, interopérabilité, et sécurité.

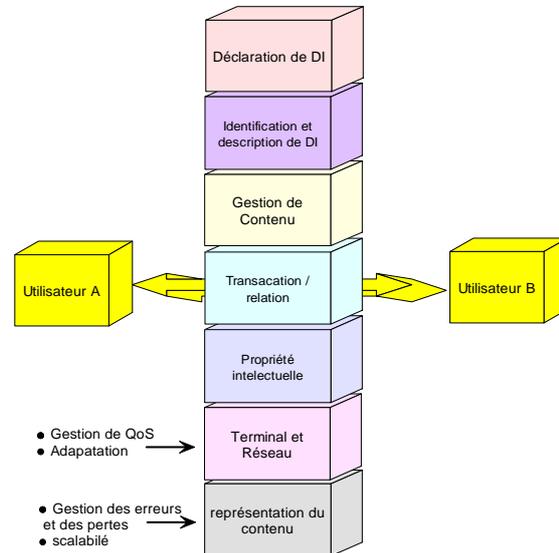


Figure 4: Les éléments architecturaux de MPEG-21

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 de deux processus : (1) processus d'adaptation de ressource, et (2) processus d'adaptation de descripteur. La Figure 5 illustre comment adapter un flux élémentaire en utilisant la machine d'adaptation définie dans MPEG-21.

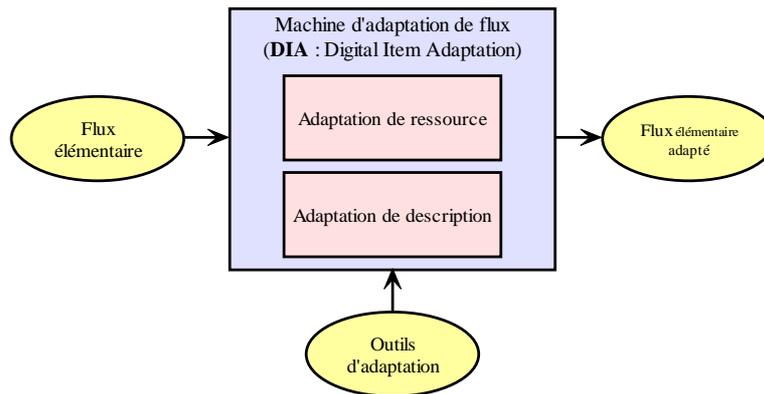


Figure 5: Machine d'adaptation dans MPEG-21

1.4. Distribution Adaptative des Paquets Vidéo sur IP avec une approche inter-couche (cross-layer)

Dans cette section nous présentons une approche d'adaptation des flux audiovisuels transportés sur IP basé sur une approche intercouche ou « crosslayer » [24]. En effet, cette approche propose une continuité de qualité de service à travers le support de la QoS sur les différentes couches du modèle OSI (ou TCP/IP). Sur cette base, nous avons proposés un système inter couche pour maintenir la QoS d'un service multimédia au niveau des différentes couches réseau. Au niveau de la couche application, les services multimédia seront adaptatifs aux changements des conditions réseau et aux variations de la bande passante. Au niveau de la couche transport et réseau, les mécanismes de réservation ou de différenciation de service seront employés pour soutenir l'approvisionnement de bout à bout de la QoS. Notamment, la couche réseau doit fournir des mécanismes de gestion de mobilité et du routage avec support de QoS. Au niveau de la couche liaison et accès au support l'ensembles des mécanismes définis dans les couches supérieurs doivent être soutenus ainsi qu'au niveau de la couche physique. Des mécanismes de contrôle et de corrections d'erreurs peuvent être employés à ce niveau pour accroître l'efficacité de la transmission multimédia.

En effet, tandis que la plupart des travaux de recherche dans ce domaine respectent le paradigme d'isolation et d'indépendance des couches protocolaires hérité du modèle de référence OSI, notre approche conceptuelle supprime 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.

Conformément aux modèles architecturaux MPEG-4 et IP Diffserv, les propositions définies dans [24] s'articule 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 avec contrôle d'erreurs et de débits ; et (3) d'un algorithme de marquage dynamique de paquets vidéo IP asservi aux variations de bande passante et aux caractéristiques des flux multimédia. L'architecture du système de distribution adaptatif des paquets vidéo sur IP est présentée dans la Figure 6.

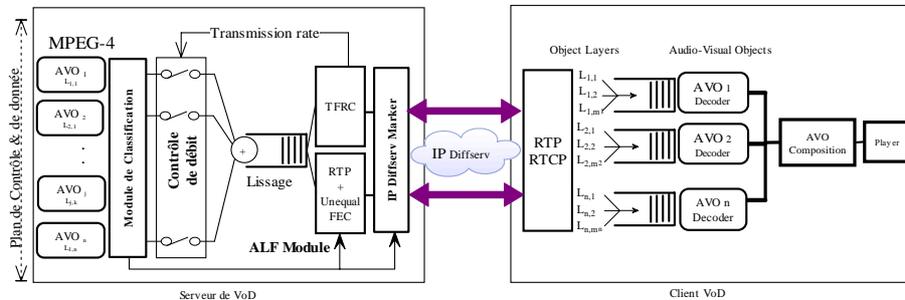


Figure 6: Architecture du système adaptatif du transport des paquets vidéo sur IP

Ce système est composé d'un serveur et d'un client vidéo interconnecté via un réseau IP utilisant le mécanisme Diffserv. Le module de classification des objets vidéo, au niveau du serveur, permet d'affecter un degré de priorité à chaque flux élémentaire de la scène. En fonction de la priorité relative de l'objet, le module de contrôle de débit régularise son débit en arrêtant la transmission des objets les moins importants en cas de détection de congestion. Le module de contrôle d'erreur utilise les codes correcteurs dynamique FEC (Forwarding Error Correction). Ainsi, les objets vidéo les plus important seront mieux protégés contre les erreurs par rapport aux objets les moins importants. Au niveau du player vidéo, les objets reçus seront bufférisés et synchronisés avec la scène et joués au bon moment.

L'ensemble des éléments qui constitue cette architecture sont résumés dans ce qui suit (voir [24][24][25] pour plus de détails) :

- un modèle de classification automatique d'objets audiovisuels intégré à la couche système MPEG-4 : ce système permet d'assurer la mappage de la qualité de service du niveau applicatif à celle du niveau réseau (système cross-layer). L'algorithme est basé sur l'utilisation des descripteurs d'objets audio/vidéo MPEG-4, et leur classification de manière automatique. Cela permet d'affecter une priorité relative à tous les objets présents dans la scène MPEG-4.
- un protocole robuste de fragmentation, d'encapsulation et de transport de flux vidéo avec contrôle d'erreurs et de débits. Ce protocole réalise la synchronisation, la fragmentation, l'encapsulation et le multiplexage optimal des flux élémentaires MPEG-4 sur les réseaux IP, réduisant 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. Ainsi, les objets

MPEG-4 exigeant les mêmes critères de QoS, seront regroupés ensemble, encapsuler sur RTP et protéger contre les erreurs en fonction de leur indice de priorité.

- Un algorithme de contrôle de débit compatible TCP (TCP-Friendly). Cet algorithme permet d'ajuster finement le débit du serveur vidéo en fonction du débit autorisé pour sa connexion. L'ajustement est contrôlé en ajoutant/supprimant des objets MPEG-4 en fonction de leurs priorités initialement calculées par le modèle de classification.

Dans ce qui suit, nous présentons l'algorithme de contrôle de débit pour l'adaptation des services audiovisuels sur IP. Cet algorithme rend le serveur vidéo au courant des conditions réseau. Il lui donne la possibilité de réagir face aux changements de ces conditions en appliquant une adaptation audiovisuelle à granularité très fine basée sur le concept d'objet multimédia introduit dans la norme MPEG-4.

1.4.1. Algorithme de Contrôle de Débit

L'algorithme de contrôle de débit est fondé sur l'utilisation d'un transport de type TCP-Friendly. En effet le mécanisme TCP-Friendly permet d'émuler le fonctionnement de TCP sur une autre couche de transport tel que UDP sans toute fois le dupliquer. Plusieurs algorithmes TCP-Friendly ont été développés dans le cadre du transport de la vidéo sur IP. L'IETF a récemment proposé l'algorithme TFRC [13] qui semble être le meilleur protocole pour assurer l'équité de trafic entre plusieurs flux. Cet algorithme de contrôle de débit est basé sur le module TFRC. Il opère comme suit :

- Le client mesure les pertes rencontrées et retourne ces informations au serveur. Cela est possible grâce au rapport de contrôle RTCP. Puisque chaque paquet RTP contient dans son entête des informations telles que le timestamp et le numéro de séquence, il est donc possible de calculer les valeurs du taux de pertes et du RTT (Round-Trip-Time).
- Les valeurs du taux de perte et du RTT sont utilisées par le module TFRC pour calculer le débit autorisé (voir Eq.1).
- Le serveur ajuste son débit en fonction du débit calculé en ajoutant/supprimant des objets audiovisuels dans la scène en fonction de leur importance. Les objets audiovisuels les moins important seront supprimer en cas de congestion.

Le débit calculé est obtenu par la formule décrite dans l'Eq. 1.

Titre de l'ouvrage en Times NR, corps 9

Dans ce qui suit, nous exposons la technique d'adaptation basée sur les objets audiovisuels. Soit E un semble d'objets audiovisuels triés par rapport à leur importance dans la scène par exemple en utilisant le modèle de classification présenté en détail dans la référence [26].

$E = \{L_{1,1}, L_{1,2} \dots L_{1,m_1}, L_{2,1}, L_{2,2} \dots L_{2,m_2}, \dots, L_{n,1}, L_{n,2} \dots L_{n,m_n}\}$, avec L , un objet ou une couche audiovisuel, cette ensemble est noté :

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

Le serveur vidéo ajoute un objet audiovisuel lorsque le débit estimé R_{TCP} excède le débit actuel du serveur. Supposons que le serveur est en train d'envoyer k entités (objets ou couches audiovisuels) à l'instant t_i , nous nous intéressons à ce qui se passera à l'instant t_{i+1} . En effet, à cet instant, le serveur doit obéir à la règle suivante (Eq.2)

$$\sum_{j=1}^{k+1} R_{i+1}(e_j) \leq R_{TCP} \quad (\text{Eq. 2})$$

Au niveau du client, les nouveaux objets audiovisuels reçus seront bufférisés et synchronisés entre eux pour être joués.

De la même manière, si le débit estimé RTCP indique que le serveur consomme plus qu'il en faut, alors il doit réduire son débit en arrêtant le streaming des objets les moins importants, et ceci tant que la règle suivante est vrai (Eq.3)

$$\sum_{j=1}^k R_{i+1}(e_j) > R_{TCP} \quad (\text{Eq. 3})$$

1.4.2. Gestion de la Stabilité Vidéo

Le module TFRC calcule le débit autorisé du serveur pour 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 player vidéo. Pour prévenir ce comportement, plusieurs mécanismes ont été pris en considération.

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, la moyenne EWMA est utilisés pour

répondre dynamiquement au changement des 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.

Le second mécanisme opère en collaboration avec le mécanisme précédent. Il permet d'assurer la synchronisation des images vidéo envoyées par un lissage au niveau des GOV (Group of video object plane). Le serveur vidéo est autorisé à changer son débit seulement au début de chaque GOV. La Figure 7 illustre le principe de lissage et de transmission de la vidéo pour gérer la stabilité.

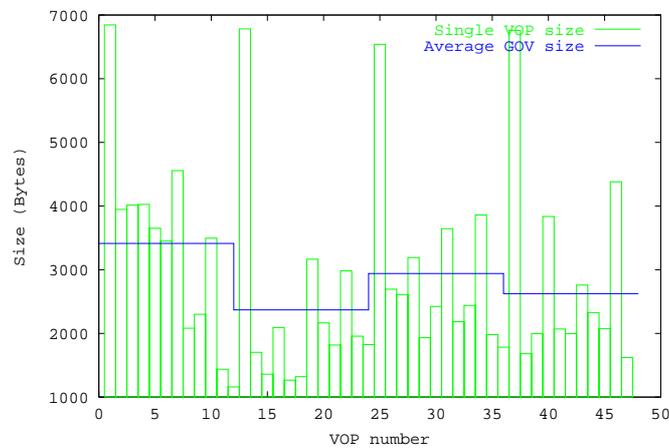


Figure 7: Gestion de la stabilité de la vidéo par lissage de GOP

1.5. Vers une distribution à grande échelle de service multimédia

Les réseaux de nouvelle génération deviennent multiservices, convergents, supportant des utilisateurs fixes, mobiles et sans-fil à travers des infrastructures et des technologies d'accès hétérogènes. La convergence des services de télédiffusion «broadcast», Internet et des services mobiles, ainsi que la transition de la TV fixe vers la TV mobile, adossé au déploiement des systèmes multimédia, représente actuellement un des objectifs primordiaux du marché de l'audiovisuel et des télécommunications. Cependant, cet objectif est loin d'être réalisé. Nous présentons dans ce qui suit quelques avancées techniques permettant d'accélérer la distribution à grande échelle de service multimédia.

1.5.1 Convergence de services broadcast, Internet et mobiles

L'intégration des réseaux de télédiffusion «broadcast» vers les mobiles DVB-H, avec les réseaux IP (Internet/Intranet) et télécoms 2.5/3G représente la prochaine étape et le grand défi technologique à relever pour la réalisation d'une convergence totale des services multimédia. En effet, les réseaux DVB offrent un service de type 1 à n avec une QoS optimale, cependant ils souffrent de certaines limites telles que la possibilité d'interactivité naturellement disponible dans les réseaux IP. Un rapprochement entre les réseaux DVB et IP est nécessaire pour offrir des services à valeurs ajoutées.

1.5.1 Interconnexion IP et DVB

L'efficacité d'interconnexion des réseaux IP et DVB repose sur le format d'encapsulation utilisé. Au sein de l'IETF, le groupe de travail IP over DVB (IPDVB) s'intéresse à la spécification des nouveaux protocoles pour accélérer le déploiement de IP sur DVB et ATSC, qui utilisent les normes MPEG-2, MPEG-4 d'une part et IPv4 et IPv6 d'autre part. Ce groupe de travail s'intéresse donc au développement de nouveaux protocoles et architecture permettant l'interconnexion avec les réseaux IP. Cependant, des verrous restent à lever concernant des problématiques bien identifiées tels que la négociation des options dans l'entête d'encapsulation, la résolution dynamique des adresses unicast pour IPv4 et IPv6, ainsi que les mécanismes nécessaires pour mapper les adresses multicast sur les réseaux MPEG-2/MPEG-4.

Le groupe IPDVB définit actuellement le protocole ULE (Ultra Lightweight Encapsulation). C'est un format d'encapsulation des paquets IPv4/IPv6 sur la couche de transport ISO MPEG-2 TS. Ce protocole décrit un format commun de base et un format étendu pour supporter d'autres informations d'entête à usage spécifique entre le réseau, l'émetteur et le récepteur de flux. Cette alternative d'encapsulation co-existera avec son ancêtre MPE (Multiprotocol Encapsulation) qui s'avère peut efficace. En effet, le format MPE ne supporte pas la version 6 du protocole IP, et peut augmenter inefficacement la surcharge dans les entêtes (overhead).

Cependant, les deux protocoles ne supportent pas les mécanismes de QoS (Quality of Service) actuellement définis dans les réseaux IP, tels que IntServ et DiffServ. Une extension de l'encapsulation ULE pour supporter les mécanismes de QoS reste ouverte pour des travaux futurs.

1.6. Conclusion

Dans ce chapitre nous avons exploré les challenges de la distribution des services audiovisuels sur les réseaux IP. Plusieurs paramètres de QoS ont été étudiés et analysés tels que : la gestion de bande passante, le contrôle des erreurs et des pertes et la gestion du délai du transfert et de la gigue. Une architecture de transport adaptatif des services audiovisuels sur IP est aussi présentée. Cependant pour permettre une distribution à grande échelle de services multimédia un rapprochement entre le monde de broadcast, IP et cellulaire est sans doute nécessaire. Dans ce contexte, l'intégration des réseaux de télédiffusion «broadcast» vers les mobiles DVB-H, avec les réseaux IP (Internet/Intranet) et télécoms 2.5/3G représente la prochaine étape et le grand défi technologique à relever pour la réalisation d'une convergence totale des services multimédia.

1.7. Références

- [1] X. Sun, S. Li, F. Wu, G. Shen, and W. Gao, "Efficient and flexible drift-free video bitstream switching at predictive frames", ICME 2002.
- [2] X. Sun, F. Wu, S. Li, W. Gao, and Y. Q. Zhang, "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.
- [3] 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.
- [4] S. Wee, J. Apostolopoulos and N. Feamster, "Field-to-Frame Transcoding with Temporal and Spatial Downsampling," IEEE International Conference on Image Processing, October 1999.
- [5] E. Amir, S. McCanne, and H. Zhang "An Application Level Video Gateway", In Proc. ACM Multimedia 1995.
- [6] E. Amir, S. McCanne, and R. Katz "An Active Service Framework and its application to real-time multimedia transcoding", in SIGCOMM, symposium on communications architectures and protocols, Septembre 1998.
- [7] S. Wee, J. Apostolopoulos and N. Feamster, "Field-to-Frame Transcoding with Temporal and Spatial Downsampling," IEEE International Conference on Image Processing, October 1999.
- [8] V. Jacobson, "Congestion Avoidance and Control", ACM SIGCOMM, August 1988.
- [9] M. Mahdavi, and S. Floyd, "TCP-Friendly Unicast Rate-Based Flow Control", Technical note sent to the end2end-interest mailing list, January 8, 1997.

- [10] S. Floyd, M. Handley, J. Padhye, and J. Widmer "Equation-based congestion control for unicast applications", In Proc. of ACM SIGCOMM, pp. 43–56, 2000.
- [11] 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" (NOSSDAV), Cambridge, UK, 8-10, July 1998.
- [12] R. Rejaie, M. Handley, and D. Estrin, "RAP An End-to-End Congestion Control Mechanism for Realtime Streams in the Internet". In Proc. IEEE Infocom '99, New York, NY, March 1999
- [13] M. Handley, S. Floyd, J. Padhye, and J. Widmer "RFC 3448, TCP Friendly Rate Control (TFRC): Protocol Specification" Request for Comments, IETF, Jan. 2003.
- [14] R. Marasli, P. D. Amer, and P. T. Conrad, "Retransmission-Based Partially Reliable Transport Service: An Analytic Model", INFOCOM, pp. 621-629, 1996.
- [15] X. Li, S. Paul, and M. Ammar "Layered Video Multicast with Retransmission (LVMR): Hierarchical Rate Control Schemes", In Proc. of Infocom 98, San Francisco, CA, March 1998.
- [16] S. Lin, D. J. Costello, "Error Control Coding: Fundamentals and Applications", Prentice Hall, 1983.
- [17] I. Rhee and S. Joshi, "Error recovery for interactive video transmission over the Internet", IEEE J. Selected Area Comm., vol. 18, no. 6, pp. 1033-49, June 2000
- [18] G. Liebl, M. Wagner, J. Pandel and W. Weng "An RTP Payload Format for Erasure-Resilient Transmission of Progressive Multimedia Streams" draft-ietf-avt-uxp-04.txt, work in progress, expires: May 2003.
- [19] 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.
- [20] K. Fujimoto, S. Ata, and M. Murata, "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.
- [21] K. Fujimoto, S. Ata, and M. Murata, "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.

- [22] E. G. Steinbach, N. Farber, and B. Girod, "Adaptive Playout for Low Latency Video Streaming," Proc. International Conference on Image Processing (ICIP-01), Oct. 2001.
- [23] M. Kalman, E. Steinbach, and B. Girod, "Adaptive playout for real-time media streaming, in Proc. IEEE Int. Symp. on Circ. and Syst. May 2002.
- [24] Toufik Ahmed, "Adaptive Packet Video Streaming Over IP Networks: A Cross Layer Approach", Thèse de doctorat, Université de Versailles, Novembre 2003.
- [25] Toufik Ahmed, Ahmed Mehaoua, Raouf Boutaba, and Youssef Iraqi "Cognitive 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, Feb. 2005.
- [26] 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, Volume 2, pp. 1091-1096, Kemer - Antalya, Turkey, June 2003.