

Un procédé de contrôle de débit pour la transmission de vidéo MPEG-4 sur l'Internet

François TOUTAIN¹, Fabrice LE LEANNEC², Christine GUILLEMOT²

¹Ecole Nationale Supérieure des Télécommunications de Bretagne,
Campus de Rennes, 2 rue de la Châtaigneraie, BP 78, 35512 Cesson-Sévigné Cedex, France

²IRISA / INRIA Rennes,
Campus Universitaire de Beaulieu, 35042 Rennes Cedex, France

ftoutain@rennes.enst-bretagne.fr, Fabrice.Le_Leannec@irisa.fr
Christine.Guillemot@irisa.fr

Résumé – Les flux multimédia transmis à l'heure actuelle dans l'Internet sont dépourvus de mécanisme de contrôle de congestion. Ceci pose problème, car l'un des principes fondateurs des réseaux IP, en terme de gestion de trafic, donne aux applications la responsabilité de ce contrôle. En son absence, les phénomènes de congestion, induisant pertes de données et délais accrus, dégradent considérablement tous les trafics circulant dans le réseau, et peuvent conduire à un effacement de ce dernier. La présente étude se focalise sur un procédé de transmission de vidéo MPEG-4 qui intègre un contrôle de congestion adapté au flux vidéo et capable de cohabiter avec les communications téléinformatiques classiques. Elle met en évidence les avantages qui découlent d'une telle approche, non seulement pour le réseau mais aussi pour la session de transmission de vidéo, qui voit la qualité finale du flux accrue.

Abstract – Nowadays, multimedia streams being transmitted across the Internet do not embed congestion control mechanisms. This is a problem, because one of the founding principles of IP networks, in terms of traffic management, gives the applications the responsibility of this control. Without the latter, congestion phenomena, yielding data losses and increased delays, greatly demper all the traffics on the network, and may ultimeteley lead to network collapse. This works concentrates on a procedure for transmitting MPEG-4 video streams which embeds a congestion control mechanism dedicated to video streams, and also able to get together well with traditional telecomputing communications. It sheds the light on the benefits of using such an approach, not only for the network but also for the video session itself, whose streams quality is increased.

1. Introduction

La transmission de flux continus de données - audio, vidéo - dans l'Internet présente des contraintes qui diffèrent notablement des communications de données "classiques" (de type transfert de fichiers). En l'espèce, les médias sonores et visuels composant une application interactive s'accommodent difficilement des délais de transmission introduits par le contrôle d'erreur d'un protocole fiable, tel que TCP [5], et par contre tolèrent généralement une certaine dégradation, résultant des pertes de paquets dans le réseau. De ce fait les flux continus de données sont à l'heure actuelle transmis au moyen de protocoles (UDP, RTP [3]) démunis de contrôle de congestion. Ceci pose problème, car le contrôle de congestion, adaptant les débits des différentes sources de trafic à la capacité du réseau, est un mécanisme crucial au bon fonctionnement de ce dernier.

Le principe de base du contrôle de congestion est celui d'un processus adaptatif impliquant une source de données et un récepteur, qui permet de réguler le débit de la source en fonction de l'évolution de l'état du réseau. Cet état est observé de manière indirecte, par la mesure de paramètres de la communication tels que le délai d'aller-retour et le taux de pertes de paquets. Dans le réseau Internet, le protocole TCP met en œuvre une telle observation, et régule le débit de la

source de données par le biais de mécanismes algorithmiques complexes (*slow-start*, évolution par croissance linéaire / décroissance exponentielle de la *fenêtre de congestion*...[5]). L'utilisation quasi-universelle de ce protocole dans le réseau est à la base même d'un fonctionnement pérenne : les congestions survenant en un point donné se résorbent d'elles-mêmes, de sorte que les délais de transmission tout comme le taux de perte de paquets sont maintenus à des niveaux raisonnables.

A l'évidence, les communications impliquant des transferts de données audio / vidéo se doivent de respecter le principe du contrôle de congestion, c'est-à-dire en fait de participer au maintien de bonnes conditions de transmission. L'objectif de cette étude est d'introduire un procédé de contrôle de débit qui, en régulant une source vidéo MPEG-4, permet au flux vidéo de cohabiter dans le réseau avec les communications TCP, tout en observant des caractéristiques de transmission adaptées à un flux continu.

2. Codec vidéo MPEG-4

Le codeur vidéo utilisé dans cette étude, présenté dans [1], vise à établir un compromis satisfaisant entre efficacité de compression et résistance aux erreurs canal (pertes de

paquets). Les techniques d'encodage classiques (codages différentiel, entropique, à longueur variable...) sont en effet très sensibles aux pertes de paquets, provoquant des erreurs de décodage qui peuvent subsister à l'intérieur d'une image jusqu'au point de resynchronisation suivant, voire s'étendre sur plusieurs images lorsqu'une prédiction temporelle est mise en œuvre. Plusieurs techniques visant à limiter l'impact de ces pertes de paquets ont été récemment étudiées, depuis les codes VLC réversibles et l'introduction dans le flux binaire d'éléments syntaxiques (points de resynchronisation, marqueurs de *data-partitioning*), jusqu'à la restriction de la fenêtre de prédiction temporelle visant à limiter la propagation d'erreurs, voire l'utilisation du rafraîchissement conditionnel de macroblocs (*conditional replenishment* ou CR) en lieu et place du mécanisme de prédiction temporelle. Le CR consiste à coder en INTRA les macroblocs de pixels 16*16 dont le mouvement mesuré dépasse un seuil fixé.

Le codeur utilisé suppose que le réseau est modélisé comme un canal d'Elliott-Gilbert (chaîne de Markov à deux états, "perte" et "non perte"), et met en œuvre une sélection de modes de codage INTRA/INTER optimisée au sens débit-distorsion. La métrique de distorsion utilisée incorpore distorsions de source et de canal, en exploitant le modèle d'Elliott-Gilbert, avec des probabilités de transitions résultant d'une observation préalable [2]. Cette sélection fournit pour chaque groupe de macroblocs un ensemble de modes de codage, de façon à minimiser la métrique de distorsion sous la contrainte d'un budget de débit. Les expérimentations montrent que ce codeur

- aboutit à un meilleur PSNR qu'un codage basé CR en l'absence de pertes ;
- est plus robuste qu'un codage MPEG-4 classique, et quasiment aussi performant qu'un codage CR en présence de pertes.

Le flot binaire généré par ce codeur est strictement conforme à la syntaxe vidéo MPEG-4, et peut donc être interprété par tout décodeur conforme au modèle de vérification MPEG-4.

3. Procédé de contrôle de débit

Des travaux récents ont montré que le débit en régime stationnaire d'une source TCP saturée (i.e. disposant d'un buffer de données infini) est en ordre inverse de la racine carrée du taux de pertes mesuré pour la communication [4]. Il est par conséquent possible d'exprimer, pour une communication donnée, le *débit équivalent TCP* que doit adopter cette communication pour se comporter de façon analogue à une connexion TCP. Plus précisément, Madhavi et Floyd proposent dans [4] la relation :

$$D_{TCP} = 1,22 \times \frac{MTU}{RTT \times \sqrt{T}} \quad (1)$$

pour approcher le débit D_{TCP} d'une source TCP dont la transmission se caractérise par un taux de perte T , un délai d'aller-retour RTT , et une taille de paquets MTU .

La mise en œuvre de ce modèle suppose l'observation de ses paramètres. En pratique, la taille maximum de paquet peut être fixe, ou obtenue au moyen d'un mécanisme de type *MTU discovery*. Un tel mécanisme fonctionne classiquement par essais / erreurs, en exploitant par exemple l'option de non fragmentation disponible dans le protocole de contrôle de niveau 3 des réseaux IP (le protocole ICMP). Des paquets de données de tailles de plus en plus petites sont successivement soumis au réseau. La non fragmentation requise par chacun d'eux entraîne une indication d'erreur de la part du premier routeur qui ne peut pas gérer un paquet de cette taille. Lorsque la destination est atteinte, la taille de paquet en cours de test est la MTU du lien. Les autres paramètres du modèle, i.e. le délai d'aller-retour et le taux de pertes de paquets, doivent être mesurés au récepteur. Cette mesure passe par un moyennage qui est pratiqué à l'aide d'une fenêtre temporelle glissante. Le point difficile est ici la détermination correcte de la taille de la fenêtre. Des études antérieures ont montré qu'une taille de fenêtre variable, fonction du délai d'aller-retour courant, fournit de bons résultats (en pratique, une fenêtre de largeur 30 RTT est adoptée).

L'exploitation d'un tel *modèle prédictif de débit* dans un but de régulation d'une source vidéo se heurte à plusieurs limites. En premier lieu le modèle a - à l'instar du protocole TCP - un comportement oscillatoire peu compatible avec le codage d'information vidéo [6]. En second lieu ce modèle n'est exploitable que lorsque le taux de perte observé est non nul. Dans le cas contraire le modèle fournit une prédiction de débit infinie, et une autre stratégie de contrôle de débit doit être utilisée. Pour ce faire nous étudions l'hybridation d'un modèle prédictif TCP avec un contrôle de débit basé sur l'observation du délai d'aller-retour.

Le principe qui sous-tend notre approche est de réaliser une adaptation en débit sans pertes autant qu'il est possible, tout en intégrant le modèle prédictif TCP présenté plus haut afin que l'adaptation de débit se conforme à un comportement de type TCP.

Le contrôle sans pertes exploite la mesure du débit reçu R , ainsi que l'observation du délai d'aller-retour (RTT) courant, qui est confronté à un seuil défini à l'aide d'une moyenne glissante du RTT augmentée d'une fraction de son écart-type. Il en résulte un débit de régulation D_{RTT} .

L'exploitation de ce procédé de régulation seul est possible. Elle conduit cependant à un contrôle de congestion conservatif - autrement dit plus sensible qu'un contrôle basé sur les pertes. En effet le délai dans le réseau résulte principalement du franchissement des files d'attente des routeurs, files gérées en FIFO (*First In First Out*). Les pertes de paquets se produisent quant à elles lorsque les buffers débordent. Par conséquent l'observation du délai permet une détection de congestion précoce en comparaison d'une observation des pertes de paquets.

Il est par contre possible d'hybrider les deux procédés de prédiction. Dans notre étude l'intégration des deux débits D_{TCP} et D_{RTT} est réalisée par moyennage si le débit résultant du modèle TCP est valide (non infini). Dans le cas contraire seul le débit résultant du contrôle RTT est pris en compte.

Les deux évaluations de débit (débit équivalent TCP et débit résultant du contrôle basé sur le délai) ainsi que leur hybridation sont effectuées par le récepteur du flux vidéo. Cette opération est menée périodiquement, et son résultat transmis à l'étage de régulation de l'encodeur vidéo. Le choix de la période de rafraîchissement est un point délicat, puisqu'il conditionne à la fois la dynamique du contrôle et la surcharge imposée au réseau par le trafic de contrôle. En premier lieu, cette période est nécessairement supérieure au délai d'aller-retour observé sur le chemin, sous peine d'incohérence : il faut en effet que le résultat d'une régulation soit observé par le récepteur et ainsi pris en compte dans l'évaluation ultérieure. Par ailleurs, lorsque le délai d'aller-retour est petit (cas d'un réseau métropolitain par exemple), la quantité de trafic de contrôle engendré peut être prohibitive. Dans la présente étude il a été par conséquent choisi d'assigner à la période de régulation la valeur maximum entre un délai d'une seconde et le délai d'aller-retour actuellement constaté.

Une architecture quelque peu différente est enfin envisageable, consistant pour le récepteur à retransmettre les paramètres de la communication mesurés (RTT, taux de pertes) jusqu'à l'émetteur, qui se charge, lui, des évaluations de débit. Une telle architecture présente l'avantage d'être implantable au moyen du protocole de contrôle de transmission de flux continu RTCP [3].

4. Résultats

Les figures 1 et 2 présentent les occurrences de pertes de paquets subies par une session de communication typique, sur un réseau de type WAN, entre les centres INRIA de Rennes et de Sophia-Antipolis. Le chemin emprunté par une telle session traverse les réseaux régionaux (WAN) Ouest-Recherche et R3T2 au moyen de liaisons à 2 Mbits/s. Ces réseaux sont eux-mêmes raccordés par le Réseau National d'Interconnexion (RNI), qui propose des débits s'étageant entre 2 et 34 Mbits/s. La session considérée consiste en une séquence vidéo encodée au format CIF à 10 images/s.

Sont confrontés les résultats obtenus avec un contrôle de débit de type CBR (constant bit rate, fixé à 384 Kbits/s), et ceux obtenus avec le contrôle adaptatif proposé plus haut. Le débit initial choisi dans ce second cas est également de 384 Kbits/s. Il est rapidement corrigé par le contrôle de congestion, en fonction des conditions de transmissions observées (pertes), pour s'établir autour d'une moyenne de l'ordre de 190 Kbits/s (figure 3).

La figure 4 présente les ratios signal-bruit (PSNR) résultant des deux scénarios précédents, i.e. transmission à débit constant ou transmission avec contrôle de congestion. L'impact des pertes de paquets sur la qualité finale du flux décodé est ici mis en évidence : en effet, l'utilisation du contrôle de congestion, bien que provoquant une diminution de moitié du débit exploitable, permet néanmoins d'obtenir un PSNR sensiblement meilleur. Ceci provient de la forte diminution des occurrences de pertes engendrée par le contrôle de congestion.

À l'évidence l'utilisation d'un contrôle de congestion permet d'adapter la communication aux conditions de réseau, ce qui a pour effet de réduire les pertes subies. Cependant ceci se produit aux dépens du débit obtenu par la source, puisque celui-ci est considérablement réduit par le jeu du contrôle hybride RTT/TCP (figure 3).

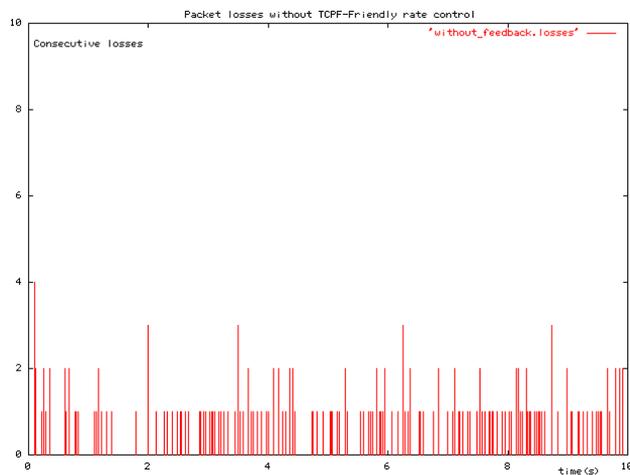
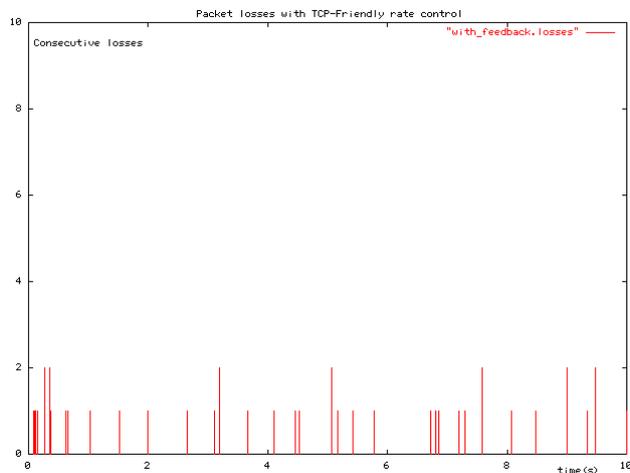


FIG. 1 : Occurrences de pertes – sans contrôle de congestion



° FIG. 2 : Occurrences de pertes – avec contrôle de congestion

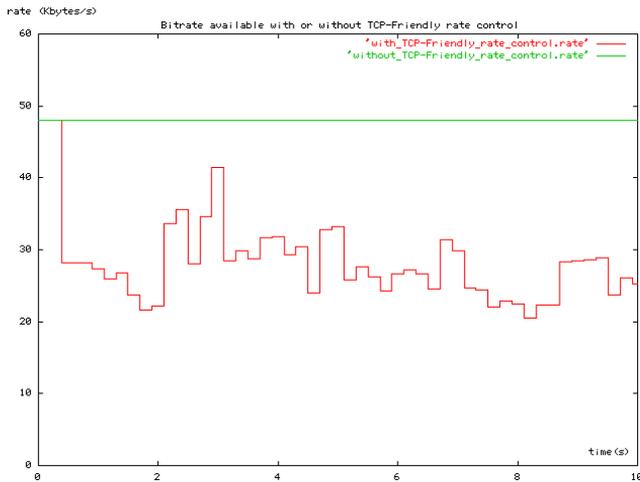


FIG. 3 : Débit imposé à la source par le contrôle de congestion

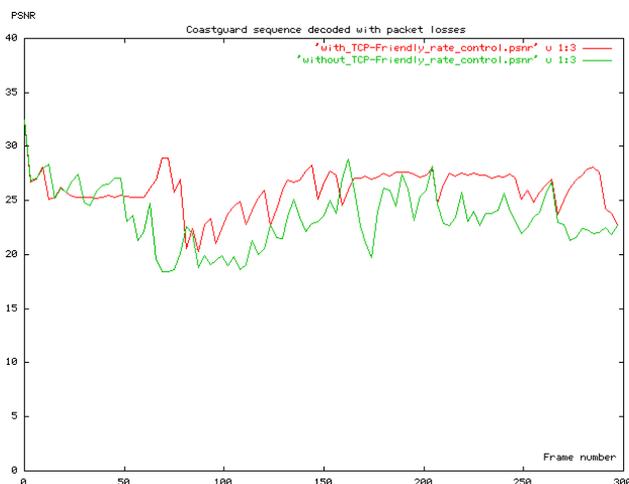


FIG. 4 : PSNR résultants, avec et sans contrôle de congestion

5. Conclusion

° ° Nous avons proposé dans cette étude l'usage de modèles prédictifs de débit pour réguler une source vidéo MPEG-4, dans le but d'une application vidéo ayant un comportement « responsable » vis-à-vis de l'utilisation des ressources de communication. Le problème visé ici est crucial du point de vue des réseaux à commutation de paquets (Internet par exemple), puisque la philosophie sous-jacente aux réseaux IP requiert de la part des applications un comportement coopératif pour résorber les congestions, et ainsi contribuer à maintenir une qualité de service acceptable. L'on note que ce problème est largement ignoré par l'industrie, qui préconise par exemple l'usage des normes ITU H.323 [7] pour le déploiement d'applications multimédia sur Internet. La suite de normes H.323 explicite en effet l'utilisation de la pile RTP / UDP / IP pour tous les flux de données impliqués dans une application multimédia (audio, vidéo), ce qui exclut *a priori* toute idée d'un contrôle de congestion.

L'on remarque également que le déploiement de technologies réseau de type « qualité de service », telles que la réservation de ressources ou la différenciation de classes mises à l'étude par l'IETF, ne résout pas le problème : il est en effet probable que ces diverses classes se distinguent par des tarifs d'accès échelonnés, et partant, que les classes offrant une qualité de service élevée soient aussi les plus onéreuses. De ce fait l'on peut prévoir qu'un certain nombre d'utilisateurs cherchent à utiliser les classes à qualité de service moindre pour transmettre des flux requérant au départ une qualité de service supérieure (il est de plus extrêmement compliqué d'interdire l'accès de certaines classes à certains flux, en se basant sur leur contenu). La problématique du contrôle de congestion, liée à la cohabitation de flux issus d'applications variées, reste donc d'actualité.

Enfin, dans une optique similaire, il est à noter que les communications multicasts (une ou plusieurs sources vers plusieurs destinations) constituent un sujet d'étude d'autant plus intéressant qu'il permet de répondre à des problèmes d'hétérogénéité (des conditions de transmission dans le réseau, des capacités de décodage et de traitement dans les récepteurs...). Il s'agit dans ce contexte de concevoir un ensemble de mécanismes permettant à une source de transmettre un même flux vidéo à plusieurs récepteurs, tout en ayant la possibilité d'adapter le débit reçu par chacun aux conditions de transmission qu'il perçoit localement. Le codage hiérarchique, associé à une transmission différenciée des couches, constitue un préalable prometteur. Notre approche peut alors être transposée à chaque sous-flux constituant la session vidéo, de sorte à protéger et réguler chacune de façon indépendante.

Références

- [1] F. Le Léannec, C. Guillemot. *Error Resilient Video Transmission over the Internet*. Procs. SPIE VCIP'99, january 1999.
- [2] J.C. Bolot, T. Turlitti. *A Rate Control Mechanism for Packet Video in the Internet*. In Procs. IEEE Infocom'94, vol 3, pp 1216-1223, Toronto, Canada, june 1994.
- [3] H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson. *RTP: A Transport Protocol for Real-Time Applications*, Request for Comments 1889, Internet Engineering Task Force, january 1996.
- [4] J. Mahdavi and S. Floyd. *TCP-friendly Unicast Rate-Based Flow Control*. Technical note sent to the end2end-interest mailing list, january 8th, 1997.
- [5] W.R. Stevens. *TCP-IP Illustrated*, Volume 1 - The Protocols. Addison-Wesley Professional Computing Series, 1994.
- [6] F. Toutain. *TCP-friendly Point-to-point Video-like Source Rate Control*. Submitted to IEEE ICMCS'99 (Packet Video'99), dec. 1998.
- [7] ITU-T *Recommendation H.323. Packet-based multimedia communications systems* (version 2). fev. 1998