

Multi-diffusion et répartition de contenus vidéos dans un réseau actif: un modèle en couches

Claude DUVALLET*, Ronan KERYELL**, Sylvain GUÉRIN**, Yerom-David BROMBERG**, Gérard BABONNEAU***

*LIH - Université du Havre
25 rue Philippe Lebon - BP 540 - 76058 Le Havre CEDEX, France
Claude.Duvallet@free.fr

**GET - ENST Bretagne
Technopôle de Brest Iroise, CS 83818 - 29238 Brest CEDEX, France
{Ronan.Keryell,Sylvain.Guerin}@enst-bretagne.fr

*** France Télécom R&D
4 rue du Clos Courtel, BP 59 - 35512 Cesson Sevigné CEDEX
Gerard.Babonneau@rd.francetelecom.com

Résumé

La généralisation des réseaux hauts débit (Câble, ADSL,...) permet d'envisager une utilisation plus large d'Internet avec la transmission de flux vidéos. On peut imaginer la généralisation des serveurs de vidéo à la demande et des chaînes de télévision. Mais la multiplication des contenus multimédia transmis sur le réseau accroît aussi le risque de provoquer des saturations de la bande passante. Il est donc nécessaire d'optimiser l'utilisation de celle-ci et de limiter le trafic au strict nécessaire tout en assurant une certaine qualité de service. Dans le projet RéActiVE¹ (Réseau Actif pour la transmission de Vidéo à l'ENST Bretagne), nous avons mis en œuvre une plate-forme réseau de transmission de contenus vidéos basée sur la distribution des contenus et des moyens de calcul sur les différents nœuds du réseau, couplée à des techniques de multi-diffusion afin de factoriser les flux.

Mots-clés : diffusion de contenu multimédia, application répartie, réseau actif, cache actif.

1. Introduction

On comprend que les problèmes sous-jacents à la généralisation du transport de flux multimédia sur les réseaux actuels résident principalement dans les capacités de transport de ces réseaux. Paradoxalement, le problème le plus important risque de ne pas être au niveau de la boucle locale. Dans ce cadre, les technologies *xDSL* sont particulièrement intéressantes dans la mesure où elles fournissent (ou fourniront à court terme) à l'utilisateur un débit conséquent lui permettant de suivre une voire plusieurs chaînes de télévision simultanément et à une qualité satisfaisante. Le développement de tels services risque bien au contraire de générer des problèmes de saturation et d'engorgement des réseaux intermédiaires, entre la boucle locale et les backbones², voire au niveau des backbones eux-mêmes. En effet, l'intérêt de déployer de telles technologies—pour se démarquer du réseau classique de radiodiffusion (et bientôt du réseau numérique hertzien)—réside bel et bien dans la possibilité de fournir à l'utilisateur l'interactivité, les applications et les services à haute valeur ajoutée qui ont fait le succès d'Internet.

Un service de vidéo à la demande sur Internet n'a d'intérêt (concurrentiellement face à la télévision classique, si on pense à la télévision numérique et les multiples bouquets de chaînes) que si on fournit à

¹ Projet financé dans le cadre du CRE établi avec France Télécom R&D et par la fondation Louis LE PRINCE RINGUET.

² Cœur de réseau central très rapide qui fédère une multitude de petits réseaux.

l'utilisateur la possibilité de pouvoir faire des pauses et se déplacer dans le flux. Ceci implique d'associer un flux à un et un seul client. La généralisation du transport de flux vidéo sur Internet nécessiterait à terme de pouvoir disposer—y compris sur les backbones—de bande passante très importante. Il est donc nécessaire de réduire l'utilisation de la bande passante [3].

L'objectif de cet article est de présenter l'architecture matérielle et logicielle retenue lors de la conception d'un réseau actif [7] pour la diffusion de contenus multimédia en déployant au sein du réseau lui-même des mécanismes de type pair à pair. Cette architecture réponds à un souci de tolérance aux pannes et de fiabilisation des échanges de flux multimédia.

2. Architecture matérielle

La circulation d'un contenu vidéo entre un serveur de vidéo à la demande et de multiples clients (qui ne visionnent pas le film simultanément) nécessite l'établissement de multiples canaux de transmission (un canal par client) ce qui peut encombrer la bande passante. Tous les clients demandent le même contenu vidéo mais en sont à différentes étapes de sa visualisation et doivent garder une interactivité sur le contenu demandé d'où la nécessité d'établir un canal par client.

Notre solution pour faire face aux problèmes d'engorgement de la bande passante consiste à réduire le chemin parcouru entre le client et serveur vidéo. Pour cela on place des nœuds intermédiaires dans le réseau utilisé qui auront pour tâche de stocker une partie ou la totalité du contenu demandé. Ces nœuds sont placés au plus près des clients afin que le chemin entre l'emplacement de stockage du film et le client soit réduit au strict minimum. La mise en place de ces nœuds RéActiVE au sein du réseau permettent de réduire la circulation des flux vidéos entre le serveur vidéo et les clients. Cette approche est issue du domaine des réseaux actifs [8, 7]. Pour compléter cette solution, nous avons recours à des techniques de multi-diffusion basées sur la technologie multicast [2, 4]. Cela permet de factoriser les flux multimédia qui doivent être transmis sur le backbone [1]. En effet, ainsi lorsque l'un des nœuds demande un film, il est envoyé sur un canal multicast à l'ensemble des nœuds intéressés qui peuvent ainsi récupérer le flux pour le stocker afin que leurs clients profitent de la proximité du contenu.

Plusieurs nœuds RéActiVE peuvent être disposés entre le serveur et la source vidéo et le nœud final auquel viendront se rattacher des clients (cf. figure 1(a)). Ce nœud final est un fournisseur d'accès au réseau RéActiVE qui se charge de répondre aux demandes des clients. Les nœuds RéActiVE vont ainsi avoir potentiellement trois rôles: (1) offrir la possibilité aux clients d'accéder au réseau RéActiVE, ils servent de fournisseur d'accès pour les clients, (2) un rôle central d'interconnexion avec les autres nœuds RéActiVE et de gestion des contenus disponibles sur un nœud (politique de cache, etc.), (3) un rôle d'interface avec la source vidéo (un serveur de vidéos, une chaîne de télévision, une webcam, etc.).

L'architecture du réseau RéActiVE repose donc sur un ensemble de nœuds actifs. Des sources vidéos sont rattachées à certains de ces nœuds RéActiVE et permettent d'injecter du contenu dans le réseau. Des clients peuvent se connecter aux nœuds placés stratégiquement au plus près des utilisateurs. L'ensemble des nœuds RéActiVE forme un maillage dans lequel certains ont un rôle d'interconnexion uniquement, servant à faire le lien entre plusieurs nœuds mais aussi potentiellement à stocker certains contenus. La stratégie de placement des nœuds reposent sur deux aspects: l'un consiste à placer des nœuds où se trouvent les utilisateurs et l'autre à placer des nœuds intermédiaires pour relier les premiers afin d'obtenir un maillage uniforme avec des possibilités de stockage intermédiaires. Pour choisir les contenus stockés sur les nœuds, on doit avoir recours à des algorithmes basés sur une notion d'intérêt du contenu par rapport aux utilisateurs.

3. Architecture logicielle des nœuds

3.1. Un modèle en couches

Dans le domaine des réseaux, le modèle en couches (OSI) fait référence [5]. Ce modèle en couches tire son intérêt du fait que chaque couche de protocole peut être décrite, spécifiée, implémentée et déployée à son niveau, en n'utilisant que les services décrits et spécifiés par la couche immédiatement inférieure. Si ce modèle en couche n'est pas reproductible dans le système que nous proposons, un parallèle intéressant peut être fait avec ce modèle, et une approche analogue (cf. figure 1(b)) peut être menée avec une perspective légèrement différente.

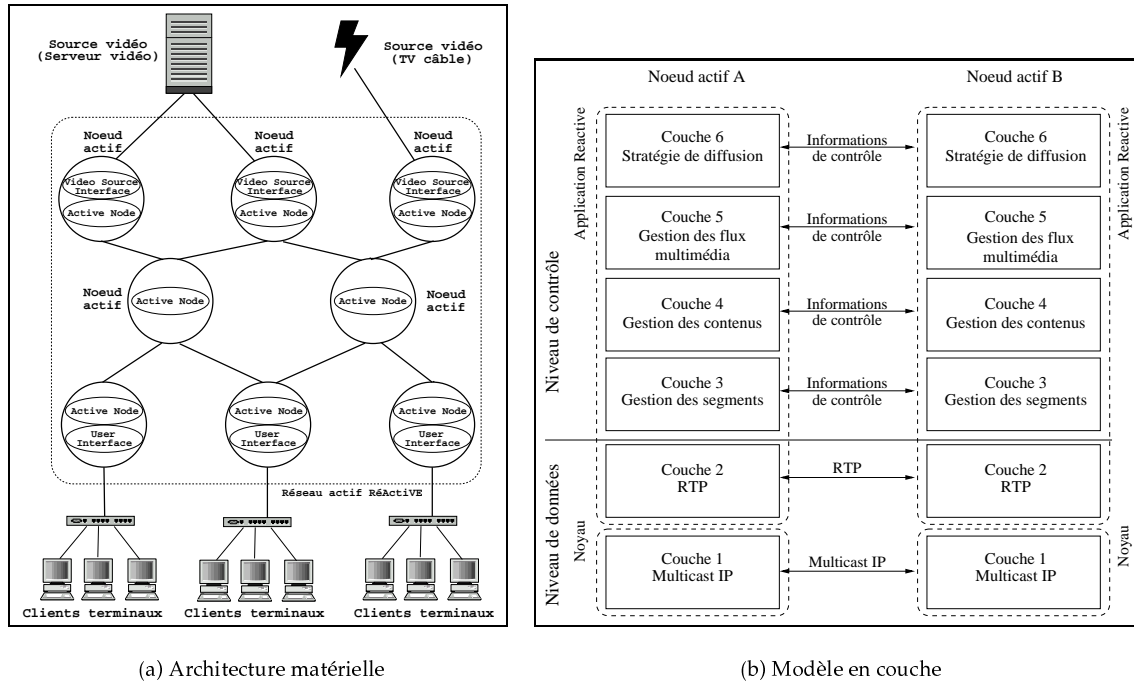


FIG. 1 – Architecture de RéActive

Un modèle en couche permet en outre de fiabiliser l'architecture du système en décomposant le problème de transmission de flux multimédia en sous-parties indépendantes.

3.2. Les différentes couches du modèle

La couche 6 permet de manipuler la notion de **stratégie de diffusion** en tenant compte de l'ensemble des contraintes du réseau dans le temps et de programmer un scénario de diffusion pour chaque requête utilisateur. Elle permet aussi de maintenir la liste—à tout moment—de l'ensemble des **stratégies de diffusion** sur tout le réseau et donc d'offrir le contexte permettant la construction de nouvelles **stratégies de diffusion** en réponse aux requêtes de nouveaux utilisateurs.

La couche 5 gère pour l'ensemble du réseau, l'ensemble des flux de contenus. Cette couche permet de maintenir la liste—à tout moment—de l'ensemble des **flux de contenus** sur tout le réseau. C'est aussi à ce niveau que sont gérés les problèmes d'allocation de canaux multicast (génération d'une adresse IP multicast et d'un numéro de port unique dans tout le réseau).

La couche 4 est une couche communicante. On peut donc la voir comme une couche distribuée sur l'ensemble des nœuds actifs, c'est-à-dire que chaque nœud coopère avec l'ensemble des autres nœuds pour remplir les services offerts par cette couche. Elle maintient à tout moment un état valide de la liste des contenus connus dans le réseau et de leur disponibilité dans chacun des nœuds connus.

La couche 3 s'appuie sur la couche de gestion des contenus et introduit la notion de segmentation des contenus. Elle offre comme principal intérêt de pouvoir représenter de façon identique et unifiée les contenus vidéos provenant de films (fichier unique par film) et ceux provenant de la numérisation en direct (par exemple à partir d'une chaîne de télévision).

La couche 2 correspond à la manipulation des **flux de données RTP** (flux de données multimédia brutes), relativement au protocole de transport adapté au cas de transport de données non-fiable, volumineuses et en temps réel : le protocole **RTP** (*Real Time Transport Protocol*) [6]. Ce protocole permet de fournir un moyen uniforme de transmettre, sur réseau IP, des données soumises à des contraintes temps réel (audio, vidéo,...). Nativement non-fiable (puisque ce protocole se base sur UDP), cette couche doit

être dotée de fonctionnalités de fiabilisation.

L'ensemble de l'architecture proposée repose sur les services de multi-diffusion (couche 1, Multicast IP). Ces services sont implantés dans le noyau du système d'exploitation.

Cette approche rend l'architecture de réseau extrêmement fiable et tolérante aux pannes logicielles dans la mesure où la chute d'un nœud actif ne provoque pas de discontinuité de service pour les flux qui n'étaient pas dépendants d'un traitement local (le système d'exploitation continue généralement de fonctionner, et de remplir ses tâches de routage et de commutation des flux multimédia).

La gestion du réseau permet de maintenir dans chaque nœud et à tout moment l'état global du réseau, des nœuds actifs en service, de la topologie du réseau, de l'activité réseau sur les connexions, etc.

4. Conclusions et perspectives

4.1. Conclusions

Notre architecture de réseau pour la diffusion de contenus multimédia sur réseau IP s'appuie sur la distribution des contenus afin de les amener au plus proche des utilisateurs et sur la multi-diffusion afin de factoriser l'utilisation de la bande passante. RéActive est avant tout une architecture logicielle (programmes actifs) reposant sur l'exploitation des infrastructures réseaux actuelles. On utilise des machines (nœuds actifs) destinées à faire fonctionner des programmes actifs et par conséquent à contrôler les flux multimédia.

Cette architecture s'appuie sur la réalisation effective d'une application sous la forme d'une plate-forme d'expérimentation. Des expérimentations ont été menées sur un réseau d'ordinateurs inter-connectés par des liaisons 100 Mbits. Il reste néanmoins nécessaire d'élargir nos expérimentations.

4.2. Perspectives

Une première piste à explorer consiste à effectuer une évaluation quantitative des performances d'un déploiement d'une telle architecture. Il conviendrait de pouvoir évaluer le gain en terme d'optimisation de l'utilisation des ressources réseau, par rapport à une architecture client/serveur classique, sur une topologie fixée (c-a-d évaluer l'économie de bande passante).

Une seconde piste à explorer consiste, sur les bases de la plate-forme logicielle actuelle, à développer de nouveaux services à plus haute valeur ajoutée pour l'utilisateur usant pleinement des capacités de calcul (grilles de programmes, mosaïque de programmes, mise à disposition de contenus par les utilisateurs, numérisation de chaînes de télévision, services de co-gardiennage, cours,...).

4.3. Remerciements

Les auteurs tiennent à remercier les personnes qui ont contribué au projet : J.A. DE LA FUENTE, F. FERRAND, Z. JANIÁK, S. JIMÉNEZ, D. MENUT, J.-F. MERCEL et B. PECCATTE.

Bibliographie

1. Asit Dan, Dinkar Sitaram, and Perwez Shahabuddin. Scheduling Policies for an On-Demand Video Server with Batching. In *ACM Multimedia*, pages 15–23, 1994.
2. S. Deering. Host Extensions for IP Multicasting. RFC 1112, Août 1989.
3. Derek L. Eager, Mary K. Vernon, and John Zahorjan. Minimizing Bandwidth Requirements for On-Demand Data Delivery. *Knowledge and Data Engineering*, 13(5):742–757, 2001.
4. K. Obraczka. Multicast transport protocols: a survey and taxonomy. *IEEE Communications Magazine*, 36(1):94–102, January 1998.
5. Pierre Rolin, Gilbert Martineau, Laurent Toutain, and Alain Leroy. *Les réseaux, principes fondamentaux*. Hermès, 1999.
6. H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson. RTP: A Transport Protocol for Real-Time Applications. RFC 1889, March 2001.
7. D. L. Tennenhouse, J. M. Smith, W. D. Sincoskie, D. J. Wetherall, and G.J. Minden. A Survey of Active Network Research. *IEEE Communications Magazine*, 35(1):80–86, 1997.
8. David L. Tennenhouse and David J. Wetherall. Towards an Active Network Architecture. *Journal of Computer Communication Review*, 26(2), 1996.