

Le Multimédia interactif sur les inforoutes.

Expérimentations : TÉLÉSIA

Alain Caristan *
Responsable du projet TÉLÉSIA
<<http://magoo.inria.fr/Telesia/>>
Tel : +33 01 39 63 56 22
email : Alain.Caristan@inria.fr

Pierre Léonard *
Tel : +33 01 39 63 50 46
email : Pierre.Leonard@inria.fr
<<http://magoo.inria.fr/pl/cv.html>>
fax : +33 01 39 63 51 88

(*)INRIA Institut National de Recherche en Informatique et Automatique
Domaine de Voluceau Rocquencourt
B.P. 105
78153 Le Chesnay Cedex

Résumé :

le débat est aujourd'hui à l'interactivité sur les réseaux informatiques. On prophétise que la société du 21ème siècle sera celle de la communication de personne à personne à travers des moyens numériques et informatiques d'interconnexion.

Depuis 1991, l'équipe TÉLÉSIA s'est intéressée à ce problème à travers deux pôles complémentaires : l'intégration/développement des nouvelles technologies dans les applications des systèmes d'informations et l'expérimentation, in situ, de ces applications à des fins de coopération interactive à travers le réseau.

Cette activité se caractérise par une dualité permanente du Subjectif et de l'Objectif. De nombreux choix techniques et organisationnels, de nombreux développements, sont le fruit d'observations pendant les expérimentations et de l'interprétation subjective qui en découle. L'évaluation et la confirmation de ces choix résultent d'une analyse quantitative qui s'appuie sur des mesures. Celles-ci nous permettent d'asseoir nos conclusions d'une façon moins arbitraire, avant d'évoluer, sans doute, vers de nouvelles orientations.

C'est pourquoi, ce papier décrit des expériences récentes menées dans le cadre de manifestations qui allient l'aspect technologique et l'aspect sociétal (coopération). La description du vécu est rapprochée des mesures réalisées dans le même temps.

Nous nous appuyons sur le dernier séminaire de l'association *Aristote*, en date du 13 Juin 1996 et sur la conférence *INET'96* des 26, 27 et 28 Juin 1996. Ces deux expériences montrent, à nos yeux, l'influence des technologies sous-jacentes sur le résultat essentiel d'une activité de coopération : la satisfaction de l'utilisateur final grâce au confort et à l'efficacité de la présence visuelle et auditive.

Nous ne prétendons pas répondre de façon catégorique à quoi servent les hauts débits ou s'ils constituent une fin en soi.

Les expériences et mesures présentées ici mettent en évidence des comportements différents de l'environnement d'expérimentation, qui ne nous semblent pas relever seulement des technologies de réseaux. Aussi, la notion de haut débit est-elle vraiment la seule à prendre en compte pour les nouvelles applications multimédia interactives ?

Faut-il introduire de nouvelles fonctions intrinsèques dans ces technologies de réseau et/ou envisager d'autres approches de la mise en application des technologies de l'information ?

1. Introduction

Depuis 1992, l'équipe TÉLÉSIA s'est fixée comme objectif de construire et d'expérimenter un service multimédia pour le travail coopératif interactif entre utilisateurs de stations de travail en réseau. L'objectif est l'échange d'information au sein d'un groupe en s'aidant sur la diffusion de l'image, du son et des données. Sur le plan technique, les différents flux d'information sont numérisés, puis compressés au moyen de logiciels spécifiques.

Le télé-séminaire est l'activité retenue pour évaluer le potentiel de faisabilité d'un service coopératif, en raison de sa simplicité sur le plan organisationnel. Il introduit cependant des contraintes déjà intéressantes. En effet, un séminaire est dirigé par un maître de conférence et suivi par des auditeurs. L'ensemble des acteurs se conforme à un protocole connu et donc plus facile à modéliser. Les fonctions à synchroniser sont déjà nombreuses et variées : sélectionner les sources sonores, les images animées, les documents du conférencier ou d'un auditeur et les aiguiller vers chacun des sites participants.

Entre 1991 et 1996, le service TÉLÉSIA a progressivement intégré les évolutions de la technologie et amélioré l'usage :

- 1991 : transmission du son et de l'image fixe (Postscript) en mode connecté et de point à point (*unicast*),
- 1992 : première retransmission TÉLÉSIA sur *RENATER* en juin 92, puis intégration d'un codec *H261* [2] logiciel et mise en œuvre de la diffusion sélective sur adresse de groupe : le *multicast*. Cette technique a permis d'optimiser la charge induite sur le réseau et d'augmenter l'impact de l'expérimentation. En effet, nous avons pu multiplier le nombre de participants, tout en simplifiant le protocole d'accès à la conférence. TÉLÉSIA intègre la gestion des flux multimédia des différents participants. La vidéo en couleur est ajoutée au début de 1993.
- 1993 : TÉLÉSIA est développé et testé régulièrement grâce à la retransmission des séminaires de l'association *Aristote*¹. Ces tests en grandeur nature, par leur régularité, et les critères de qualité qu'ils imposent, sont le contexte privilégié de la mise au point de TÉLÉSIA. Le concept de droit de parole, et sa distribution par des techniques d'asservissement à distance, devient le pivot du protocole du séminaire, car il introduit une vraie possibilité d'étudier l'impact de ces technologies sur les activités de coopération. A l'été 1993, les difficultés rencontrées dans la mise en place des liens *MBONE* conduisent l'association *ARISTOTE* à organiser le réseau *multicast* français sous le nom de *FMBONE*. Cette initiative, coordonnée par l'*INRIA*, entraîne la collaboration de nombreux sites français. Le noeud racine du réseau est installé à l'*EDF-DER* de *Clamart*.
- 1994 : le télé-séminaire *ARISTOTE* est entré dans les mœurs et il devient nécessaire de faire progresser la technique afin d'obtenir une meilleure qualité du service, qui semble souvent altéré par des problèmes de réseau. Que ce soit sur *RENATER* et le *FMBONE*, ou encore sur *ATM-Transrel* [4], TÉLÉSIA a permis d'observer les différents effets subjectifs causés par les pertes des paquets et le délai de transit. La saturation du réseau Internet, qui a une incidence directe sur le débit en K bits/s ou le nombre de paquets/s, a impliqué l'adaptation du service TÉLÉSIA à ces conditions de transmissions difficiles. Pour réduire la sensibilité des décodeurs aux pertes de paquets et améliorer la restitution de l'image, nous avons défini et implanté une nouvelle mise en paquets [12] des flux vidéo, qui tient compte de la structure du flux des informations. L'amélioration du débit de la vidéo, en terme de trames par seconde, la volonté de diffuser un son mieux défini, l'augmentation de la charge globale de l'*INTERNET* et de *RENATER* font penser qu'il y a urgence à étudier des solutions

1. L'association *Aristote* est un groupe de grands utilisateurs de moyens informatiques, répartis dans les secteurs aussi divers que l'éducation (*Ecole Polytechnique, HEC...*), la recherche (*INRIA, CNRS, CNET...*), l'industrie (*CNES, Siemens Nixdorf, Thomson-CSF...*), les services avec la *BNP*, la *Cisi, Euriware...* Rassemblés par groupes de recherche thématiques, ils joignent leurs efforts d'études prospectives, de formation, d'expérimentation et de promotion de nouveaux services.

- s'appuyant sur des réseaux de plus haut débit.
- 1996 : un système de transmission de documents : *Dia_Dif* est ajouté à TÉLÉSIA. Les transparents des conférenciers sont désormais intégrés aux présentations en temps réel. Ce système inter-opère avec les produits standards de présentation basés sur le WEB¹. TÉLÉSIA devient, ainsi, un outil de travail coopératif de l'Infrastructure Globale d'Information (Global Information Infrastructure)².

Dans ce papier, nous présentons les expériences récentes réalisées sur des infrastructures de communication expérimentales. Elles font intervenir plusieurs média de communication qui produisent des effets sur la qualité des transmissions. L'utilisation de TÉLÉSIA, qui intègre des moyens de mesure et peut donc faire de l'observation multi-sites, permet de procéder à des évaluations sur la relation entre les qualités objectives et subjectives d'une activité interactive multimédia sur réseau.

2. Principes d'expérimentations.

La plupart de nos expérimentations se déroulent sur *RENATER*. Le Réseau National d'Interconnexion est à 34 M bits/s tandis que les plaques régionales fonctionnent à des débits soit équivalents, soit plus faibles. Les points d'accès des entreprises vont de 64 K bits/s à 34 M bits/s. L'hétérogénéité des débits rend les conditions d'émission et de réception sur chacun des sites très inégales. Sur une activité de coopération en temps-réel, pour des configurations identiques de traitement, on constate aisément que les participants ne vivent pas les événements de la même façon, parce qu'ils ne disposent pas des mêmes conditions de réception. Leur coopération s'en trouve affectée.

2.1. Conditions d'expériences

Nous ne retenons que les expériences réalisées en 1996, sur des infrastructures hétérogènes, intégrant une partie de réseaux qualifiée de "haut-débit".

2.1.1. Les Séminaires ARISTOTE

L'association *Aristote* s'est donné pour objectif d'anticiper l'évolution de la société, d'identifier les innovations technologiques, d'en maîtriser la complexité et de développer les compétences qui en facilitent le transfert vers les entreprises membres. Dans le cadre du Collège de l'École *Polytechnique*, l'association propose un cycle annuel de 6 séminaires sur les réseaux et les systèmes d'information, qui sont diffusés depuis trois ans sur le réseau Internet et le *FMBONE*. Les participants sont des chercheurs, des enseignants répartis sur l'ensemble du territoire métropolitain.

La retransmission des séminaires est réalisée en coopération avec les services informatiques et audiovisuels de l'école *Polytechnique* située à *Palaiseau*. Depuis la mise en œuvre du système de diffusion *Dia_Dif* en Mars 1996, les responsables scientifiques du séminaire collectent les documents et les transparents des conférenciers dans un format électronique et les transforment en images *GIF* ou *JPEG*. Ces images sont stockées sur l'un des serveurs *WEB* de l'équipe TÉLÉSIA. Le service *Dia_Dif* permet de piloter l'affichage à partir de n'importe quel station sur le réseau en utilisant un "navigateur" *WEB*. Un ordre de contrôle *Dia_Dif* indique l'URL à charger et déclenche la requête "http" du navigateur de chaque participant. Pour les documents non électroniques, on utilise un statif de projection équipé d'une caméra.

Deux stations de travail servent à coder la vidéo des caméras et le son de la salle, à décoder les informations arrivant des participants distants et enfin, à gérer le séminaire. On utilise deux caméras, l'une pour un plan global de l'estrade, l'autre pour des plans alternés du conférencier

1. World Wide Web : système distribué de documents référencés dynamiquement par des hyperliens.

2. Politique de développement économique des télécommunications, infrastructures et services, définie par le G-7.

ou des intervenants dans la salle. Un projecteur vidéo affiche des images d'écran comportant les documents électroniques et les images des participants distants sélectionnés.

2.1.1.1. Caractéristiques des moyens de communication

Le réseau de l'École Polytechnique est maillé. Lors des premières expériences, nous utilisions le routeur *multicast* de l'École qui s'est avéré trop sollicité par les élèves qui consomment une trop grande partie de ses ressources. Les flux *multicast* qu'il devait router étaient très dégradés avant même d'arriver à la racine *FMBONE* de Clamart. Rapidement, il a été nécessaire d'installer un routeur dédié sur l'une de nos machines de retransmission. Cette simple modification contribue à une amélioration sensible de la qualité du transport.

Cette année le *FMBONE* a été doté d'une double racine : l'*EDF Clamart* et l'Université de *Jussieu*. Les deux sites sont reliés à *Renater* à 34 M bits/s. Il apparaît cependant que le routeur situé à l'*EDF* subit ou provoque localement un taux de pertes très important. Le tunnel tiré entre l'École Polytechnique et le *FMBONE* a donc été changé en faveur de celui de l'université de *Jussieu*. On constate une amélioration subjective des conditions de réception sur de nombreux sites. Il faut cependant admettre qu'il est pratiquement impossible de suivre, dans des conditions acceptables, un télé-séminaire *TÉLÉSIA* depuis une liaison à 64 K bits/s, surtout si elle est utilisée pour l'ensemble du trafic Internet de l'entreprise.

2.1.1.2. Caractéristiques des flux son et image :

- le son est monophonique. Il est codé en μ law à 8 K hz ce qui correspond à un débit de 64 K bits/s. Le débit peut-être réduit dynamiquement à 32 K bits/s en codage *ADPCM*, ou à 13,5 K bits/s en *GSM*. Cependant, ce dernier produit un timbre métallique provoquant une certaine fatigue chez l'auditeur et une perte de l'attention. De plus, le décodage *GSM* consomme de la puissance de calcul qui sature les stations réceptrices un peu faibles.

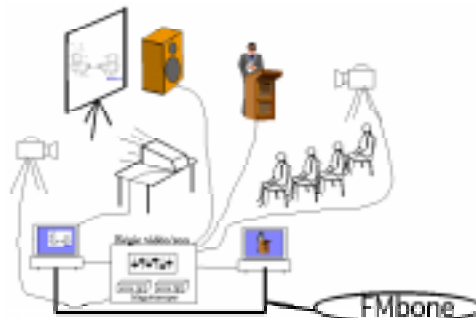


Figure 1 : Synoptique station émettrice : Ecole Polytechnique, séminaires Aristote

- les images du conférencier et de la salle sont codées en format *QCIF* (176 x 144). La fluidité peut atteindre 15 images/s pour un débit maximum compris entre 100 et 250 K bits/s,
- l'image vidéo d'un transparent est codée en *CIF* (352 x 288) et parfois en *SuperCIF* (704x576) lorsque les graphiques sont trop petits. Le débit peut alors atteindre 300 K bits/s, durant les changements de document.

2.1.2. La Conférence INET'96 à Montréal : juin 96

Dans le cadre de la conférence annuelle de *L'INTERNET Society (INET'96)*, le *GIP-RENATER* [8] et les partenaires Canadiens de *CANARIE* [1] ont décidé de démontrer l'intérêt de l'évolution vers le haut débit en réalisant une expérience de vidéo conférence transatlantique *haute qualité* entre une salle de réunion du palais des Congrès de *Montréal* et la salle audiovisuelle de l'*IUT d'Orsay*.

Le réseau haut-débit utilisait, pour la circonstance, des liens français, américains et canadiens. L'objectif français était de préfigurer la future version à haut débit, *RENATER 2*, et d'y associer des auditeurs francophones (enseignants, chercheurs, responsables) afin qu'ils puissent comprendre le bénéfice qu'eux-mêmes et leurs organismes en tireraient. Cette opération, qui a couplé, pendant plusieurs heures chaque après midi, (heure française) deux sites de référence, a été l'occasion d'initialiser un premier axe de coopération sur le co-développement. L'équipe *TÉLÉSIA* a pu, à cette occasion, analyser les nouvelles contraintes et observations qui résulteraient de l'utilisation de liaisons *IP-ATM* [5] en mode datagramme, dans un contexte où les utilisateurs se préoccupent plus des conditions finales de travail que de la prouesse technologique.

2.1.2.1. Les ateliers virtuels *INET'96*.

Les activités qui se sont succédées illustrent plusieurs modes de travail coopératif. Cela correspondait à des conditions différentes d'interaction, des volumes d'information variables et des besoins, en terme de qualité, pas toujours équivalents :

- retransmission en direct et en différé des sessions de la conférence de l'*INET'96*,
- organisation de tables rondes transocéaniques consacrées au télé-enseignement, à la stratégie de réalisation d'un technopole virtuel entre *Montréal* et l'*Ile de Science* (Plateau de *Saclay*) et aux premières réflexions sur sa mise en œuvre,
- participation à l'inauguration mondiale du site *WEB* de la chaîne de télévision Internationale *TV5*. Cette manifestation mettait en relation des personnalités politiques Québécoises et Nord Américaines à *Montréal* et des responsables Européens de *TV5* à *Orsay*.

2.1.2.2. Aspects Techniques

La plupart des caractéristiques techniques sont reprises des séminaires *ARISTOTE*. Mais, la participation d'un groupe de personnes, comme à *Orsay*, imposait de fournir à la régie audiovisuelle des images suffisamment grandes et fluides pour qu'elles puissent être projetées sur un écran géant. Le format *QCIF*, habituellement utilisé ne suffisait pas, même quadruplé à la réception. Nous avons décidé d'émettre dans le format *CIF*, puis de l'agrandir en format *SuperCif* à la réception. Le choix des dernières stations *ULTRASPARC* s'est imposé parce que leur processeur présente des caractéristiques intéressantes pour les applications de traitement du signal. Il intègre des fonctions dédiées qui accélèrent le traitement et la compression des flux vidéo et audio. Dans une application où l'ensemble du processus est à la charge du processeur, de telles fonctions apportent un gain appréciable de performance. Nous avons pu obtenir une bonne fluidité de l'image avec un débit minimum de 15 images/s environ, là où les stations classiques plafonnent à 7 images/s. En contrepartie, cela se traduit par une augmentation du volume d'information multimédia généré.

L'infrastructure initiale prévoyait une liaison de bout en bout en *IP-ATM* qui aurait impliqué le réseau de coopération Européen *ATM-PILOT*. A la suite d'un incident sur un câble transatlantique, cette configuration n'a pas pu être mise en œuvre. La liaison finalement utilisée (un chemin virtuel de 3 M bits/s sur *ATM* aboutissant à *RENATER*) était opérée par *Télglobe* sur le territoire canadien, par *Sprint* sur le territoire américain, par *Sprint* et *France Télécom* sur l'Atlantique Nord jusqu'au centre *Eurogate* et par *France Télécom* jusqu'au centre d'interconnexion avec *RENATER*. Le dernier tronçon empruntait l'épine dorsale de *RENATER* à 34 M bits/s jusqu'au centre de l'*IDRIS* du campus universitaire d'*Orsay*. Cette modification de dernière heure a influencé l'expérimentation de façon importante. Les observations et les mesures, notamment sur le délai, ne sont plus aussi déterminantes sur le plan scientifique en raison du mélange de réseau (commutation *ATM* versus routage *IP* classique).

Pour maîtriser les débits entrants et sortants, et éviter toute pollution réciproque avec le reste des activités du *MBONE*, des ponts *multicast* dédiés ont été mis en place et isolés du *FMBONE*.

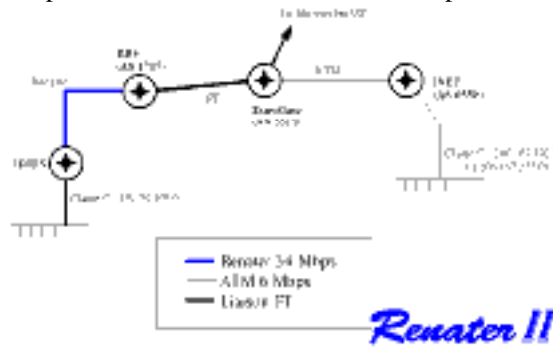


Figure 2 : INET'96 Plan d'adressage

A la mise en route, plusieurs incidents ont montré l'aspect encore très technique de cette expérimentation. Divers problèmes de serveur de nom, de ponts *multicast* instables, de tronçons peu fiables, ont démontré que non seulement la mise en œuvre, mais aussi l'efficacité, des applications distribuées temps réel dépendent du niveau de maîtrise du réseau. Cette notion de réseau maîtrisé ou contrôlé dépasse le simple problème de l'installation de lignes et de routeurs. Les réseaux intègrent désormais des services avancés nécessaires et souvent indispensables au bon fonctionnement général : domaines, sous domaines, désignation, routage, filtrage, diffusion, groupes de diffusion, attribution des numéros de groupes de diffusion...

Une fois ces problèmes réglés, les activités coopératives se déroulèrent pratiquement sans encombre, et on peut le dire, avec une qualité qui a surpris aussi bien les spécialistes que les néophytes.

3. Mesures et évaluations

Toutes ces expériences amènent des commentaires relativement subjectifs sur le rendu des images, du son et aussi sur le service obtenu au niveau de l'activité de coopération en cours. Il est encore difficile, aujourd'hui, d'évaluer scientifiquement l'efficacité des outils multimédia pour le travail de coopération. Nous avons intégré dans TĒLĒSIA des outils de mesure en temps-réel qui devraient permettre de mieux identifier et définir les paramètres de services.

Ce chapitre présente les caractéristiques de nos mesures et les premiers résultats qui en découlent. Dans l'état actuel de nos connaissances, nous souhaitons confirmer, ou infirmer, nos impressions subjectives par la mesure des conditions de réception sur chaque site. Les descriptions et analyses sont principalement basées sur une observation des phénomènes audio et visuels. Les mesures sont présentées en parallèle, sous une forme graphique. Nous savons que la qualité intrinsèque des communications influe sur le rendu, au point de restituer difficilement la synchronisation de certaines activités coopératives humaines, s'il persiste trop d'incidents de transport.

3.1. Conditions de mesures

Chaque instance active de TĒLĒSIA diffuse ses propres mesures, sur un canal *multicast* commun, qui sont collectées sur le site de *Rocquencourt*. Suite à des travaux antérieurs sur les variations de délai de transit en fonction de la mise en paquet [12], les mesures actuelles concernent, pour chaque point participant, le taux de perte de paquets et le jitter pour les flux audio, vidéo et contrôle. Il existe deux types de rapports statistiques :

- sur les pertes cumulées concernant les flux audio et vidéo. Ce rapport est mis à jour dans la

couche de gestion des protocoles *RTP* [6] propre à *TÉLÉSIA*. Il fournit des pré-analyses telles que la taille des séries de paquets perdus. Cette information peut révéler les incidents liés au routage,

- sur le jitter des flux audio et vidéo. Le calcul du jitter est défini dans les documents *RTP* [7]. En raison du manque de synchronisation fine des horloges des calculateurs et de la non monotonie de l'algorithme choisi dans *NTP* [3], il ne nous a pas été aisé de calculer le délai de transit des paquets. En conséquence, nous calculons la variation du jitter (**interarrival jitter**). Ce chiffre est obtenu en faisant disparaître les références au temps absolu des horloges par une simple soustraction des délais. Dans la suite du document, nous parlerons simplement de **jitter**.

Les pertes cumulées sont présentées sous forme d'histogrammes en trois dimensions. L'axe vertical des y correspond au nombre de fois qu'une perte de n paquets consécutifs a été constatée. L'axe des x représente la largeur des séries de pertes : c'est-à-dire le nombre de paquets consécutifs perdus en séquence. Nous nous limitons à 9 classes, la dernière comptabilise les observations égales ou supérieures à 9. L'axe transversal des z est celui du temps, découpé en tranches horaires. Le taux inscrit en marge de cet axe est, pour cette tranche, le taux de perte par rapport au trafic global enregistré. Il permet de relativiser les valeurs des barres d'histogrammes, qui sont en mesures absolues.

Le jitter est représenté sous forme d'histogrammes en trois dimensions. L'axe vertical des y représente le nombre de fois où une valeur du jitter x est rencontrée dans une classe donnée. L'axe des x donne la classification des valeurs observées pour la variation du jitter. Cette variation est comptabilisée par tranches de 1 milli-seconde pour le flux audio et par tranches de 200 micro-secondes pour le flux vidéo. Cette différence d'échelle améliore la lisibilité respective des graphiques. L'axe de profondeur z représente les tranches horaires successives.

3.2. Présentation des résultats.

Les expériences précédemment décrites ont fourni des mesures pendant plusieurs heures. Les conditions de transport et de réception ne sont pas toujours constantes, mais il a été possible d'obtenir des mesures correctes pour des activités de natures différentes. Sont présentées, ci-après, les mesures du séminaire *ARISTOTE* du 13 juin 1996 ainsi que les mesures de la conférence *INET'96* du Mercredi 26 juin et du Vendredi 28 juin.

3.2.1. Séminaire *ARISTOTE* du 13 juin 1996.

Pour cette journée du 13 juin, nous avons reçu les mesures des sites suivants : le *Cnusc* de Montpellier, l'*Irisa* de Rennes, le *CISM* de Lyon 1, l'*Ifremer* de Brest, l'*INRIA* de Sophia Antipolis, le *Sept* et l'*IN2P3* de Caen, l'université d'Evry, le laboratoire *Infobiogen* de Villejuif. Tous ces sites n'ont pas suivi avec le même intérêt les interventions de la conférence, ni décodé l'ensemble des flux générés. Nous avons pu, néanmoins, déterminer une plage horaire entre 9 et 13 heures qui rassemble le maximum de sites ayant "écouté" et "regardé". Ces sites sont : le *Cnusc*, le *Sept*, l'université d'Evry et le laboratoire *Infobiogen*. Pour des raisons de concision de ce papier, nous ne présentons ici que le site du *SEPT* de Caen, les autres documents sont disponibles sur le *WEB* [13].

Les résultats “absolus” de l’audio sont sensiblement voisins dans les trois tranches horaires. Dans la tranche de 10 à 11 heures, le taux de perte relatif est faible par rapport à la tranche de 11 à 12 heures :

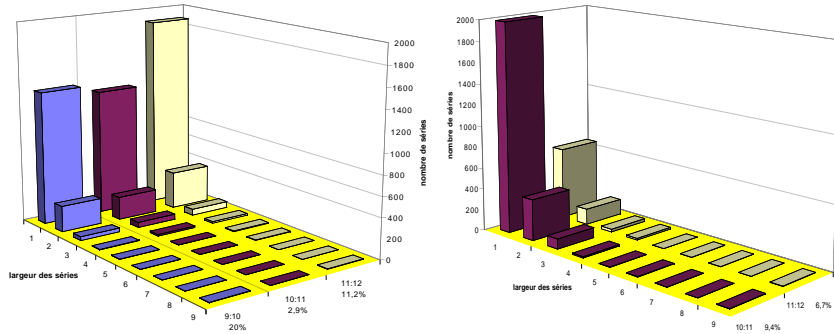


Figure 3 : Sept, mesures des pertes audio et vidéo

Deux explications sont possibles : soit une brusque amélioration des conditions de réception, soit, ce qui paraît plus vraisemblable, le participant n’a pas décodé le flux sonore pendant toute l’heure. Cela peut se comprendre si le confort auditif était dégradé, même en cas de perte d’un seul paquet en séquence. En effet, il faut aussi considérer les 20% de perte sur le total des paquets reçus comme l’indique la tranche horaire précédente.

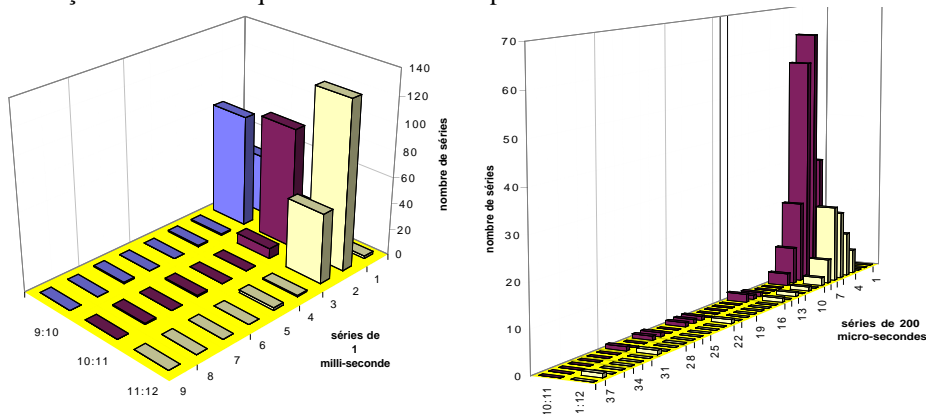


Figure 4 : Sept, mesures du jitter audio et vidéo

Le jitter audio est très régulier. Il semble se dérégler d’une demi milli-seconde au cours de la troisième heure (11 h - 12 h). Cela correspond à l’une des périodes les plus chargées de Renater. Le jitter vidéo est aussi très régulier et reste dans le créneau de 1 à 2 milli-secondes.

Ces mesures confirment le pronostic de l’influence forte du réseau. Cependant, dans l’état actuel de la technologie mise en œuvre, il est difficile de faire la différence entre les paramètres locaux de la station réceptrice et les problèmes posés par le réseau. Pour des raisons techniques d’intervention à distance sur les systèmes opératoires, il est difficile de faire des mesures en dehors de l’application. Cela signifie que certains événements ne sont détectés qu’au niveau interne de TÉLÉSIA. Les mesures dépendent donc du fonctionnement global de la station réceptrice. Néanmoins, la comparaison des commentaires du participant et les mesures ci-dessus permettent de proposer des modifications au niveau du réseau et aussi de la configuration de travail. On peut ainsi mieux évaluer la puissance nécessaire en terme de station et de réseau.

3.2.2. INET'96, Mercredi 26 Juin

On peut distinguer trois activités principales durant cette journée. Elles sont, toutes les trois, développées dans ce papier, car elles apportent, à notre avis, de précieux renseignements sur la mise au point des outils de travail coopératif lors des expériences :

- la première est le réglage du duplex audio-vidéo et la recherche des paramètres optimaux pour le rendu subjectif, c'est-à-dire la qualité du service perçue dans ses aspects audiovisuels,
- la seconde est la diffusion des conférences se déroulant dans la salle plénière d'INET'96. Cette situation demande le plus grand respect de l'image et du son de l'événement,
- la troisième est une table ronde dans une salle virtuelle *Montréal-Orsay*. Cette situation introduit les problèmes d'interaction sur de grandes distances.

D'une manière générale, c'est-à-dire, quel que soit le type d'activité en cours, les premiers tests sur le son ont révélé des délais de transit anormalement élevés (de l'ordre d'une dizaine de secondes) sur un tronçon de réseau et sur certaines machines (*ULTRASPARC*). D'autres effets passagers s'y sont ajoutés : une duplication de paquets qui produisait un son de type " **break** " (cf. les samplers en musique électronique).

L'analyse des statistiques à posteriori [figure 5] montre que seul l'audio est affecté, ce qui écarte une erreur de configuration du réseau ou une série d'accidents. L'analyse de ce phénomène est complexe et oriente vers deux causes possibles : une modification de comportement du pilote asynchrone du périphérique audio, associé à un problème d'ordonnement des processus par le noyau du système d'exploitation. La suite de la journée nous a fait suspecter le couplage entre la table de mixage du son et l'interface de la station. Le branchement direct des microphones sur l'entrée de la carte de numérisation du son, a semblé supprimer ce phénomène. Mais, le fait que cela se soit produit très souvent au début des activités journalières, nous avait orienté vers une analyse des conditions de routage *multicast*. Aucune défaillance, ni aucune saturation apparente n'a pu être détectée. Il faudrait sans doute reprendre cette hypothèse dans une analyse spécifique du routage *multicast* entre réseaux hétérogènes *IP* classique et *IP/ATM*. Finalement, la diffusion de la session plénière d'inauguration a débuté vers 14h30 et l'on n'a plus relevé d'incidents sur le plan subjectif.

Les activités de diffusion et de conférence entre *Montréal* et *Orsay* se sont ensuite enchaînées sans difficulté. Un décalage sensible entre le son et l'image subsistait cependant. Ce phénomène est connu puisqu'en l'absence de synchronisation entre les deux flux audio et vidéo, il existe un délai dans le traitement de chacun d'entre eux. Mais le phénomène a paru trop important sur les stations *ULTRASPARC*. Le taux d'occupation du processeur était de 50%, ce qui a priori permettait le décodage de l'audio et de la vidéo sans retard pour aucun des flux. Or le flux audio était soumis à une plus grande variation des délais de transit et l'analyse des mesures [figure 6] et [figure 8] va le révéler.

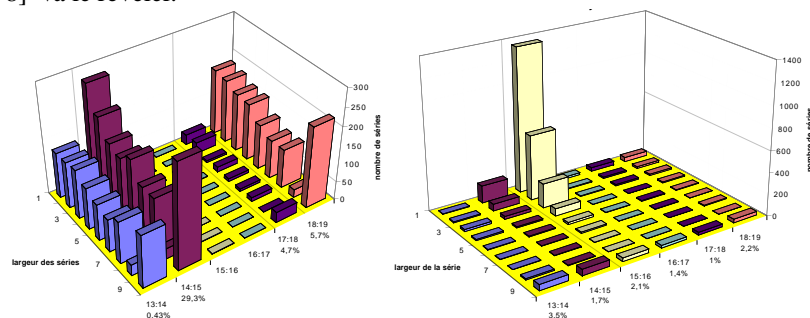


Figure 5 : Orsay 26 Juin, mesures des pertes audio et vidéo

Les pertes du site d'Orsay expliquent la dégradation du son entre 14 à 15 heures. Le taux de pertes de paquets est de 29%, avec un étalement sur l'ensemble des séries. Si on le compare aux résultats de la vidéo qui est de 1,7%, on exclut des accidents de réseau qui devrait alors affecter les deux flux. Il s'agit ici du fameux son "breaké" que l'on peut vérifier par le taux important de paquets doublés c'est-à-dire qui portent le même numéro de séquence RTP. De plus ce phénomène perturbe le système de mesure.

Les tranches horaires 15 heures et 16 heures [figure 5] correspondent à la retransmission des conférences INET'96 et le branchement en direct des sources sonores sur les stations.

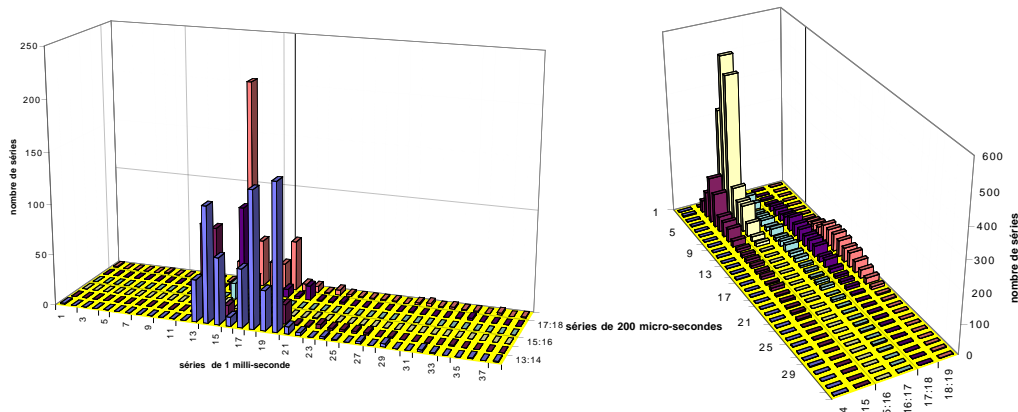


Figure 6 : Orsay 26 Juin, mesures du jitter audio et vidéo

On voit ici l'illustration du décalage des délais entre les flux audio et vidéo. Le jitter du flux audio est environ de 17 milli-secondes alors que celui du flux vidéo est dans une fourchette de 0,2 à 4 milli-secondes, avec une augmentation en fin de journée.

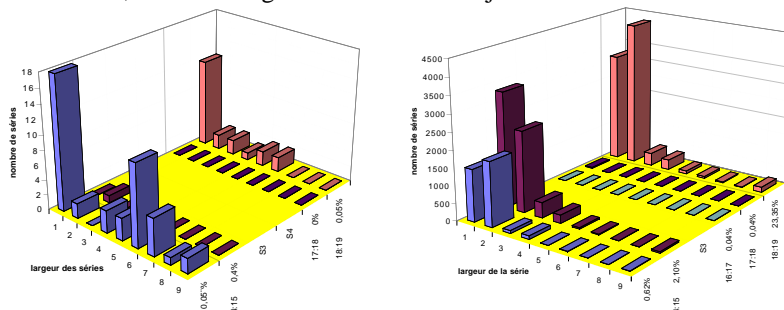


Figure 7 : Montréal 26 Juin, mesures des pertes audio et vidéo

Les pertes de l'audio sur le site de Montréal sont régulières. Les tranches horaires vides correspondent, soit aux retransmissions des sessions de la conférence, qui ne nécessitent pas de décodage dans le sens Orsay > Montréal, soit à une période exempte de perte. Par contre, la vidéo reçue à Montréal est proportionnellement plus affectée que sur le site d'Orsay. Une hypothèse est avancée pour expliquer cette dissymétrie :

- les tronçons ATM allant de Renater au palais des Congrès ont été négociés, réglés et testés séparément sans vision globale du service attendu. Il était par ailleurs difficile de prévoir les conditions réelles de fonctionnement,
- il existait un effet d'entonnoir au centre Eurogate d'interconnexion entre le flux de Renater à 34 M bits/s et la liaison ATM à 3 M bits/s. Dans le sens Orsay > Montréal on obtenait une

surcharge du routage vers *ATM* qui provoquait une perte de paquets. Cela affectait de façon visible l'image reçue à *Montréal*. L'utilisation simultanée de différentes machines en réception, l'observation de leur charge qui restait très correcte, alors que les effets observés étaient identiques, semble valider sérieusement cette interprétation.

Les conditions d'interconnexions des réseaux sont encore mal connues ou mal estimées. L'accent est mis sur la bande passante en M bits/s alors que les applications interactives requièrent aussi un maintien du débit en paquets/s. C'est-à-dire une puissance de commutation suffisante qui nécessite un processus de routage, de filtrage plus rapide. Le redémarrage des pertes lors de la dernière tranche de 18 à 19 heures provient de la reprise des tests en préparation de la seconde journée. Nous recherchions un certain nombre de réglages des paramètres audio et vidéo permettant de déterminer les conditions limites de fluidité de l'image et du son.

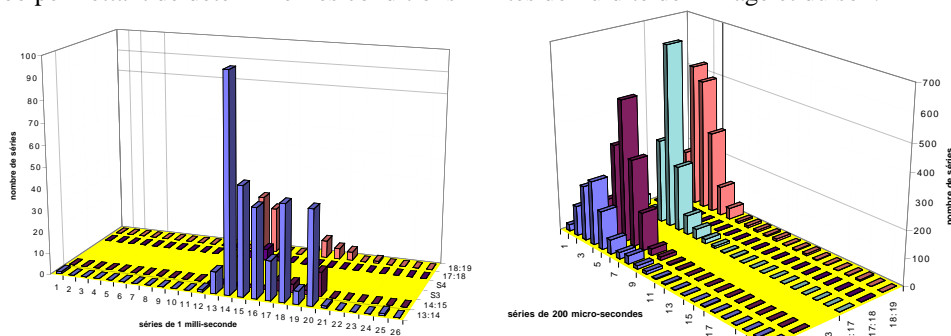


Figure 8 : Montréal 26 Juin, mesures du jitter audio et vidéo

A *Montréal*, le phénomène de décalage du jitter audio par rapport à celui de la vidéo est identique. L'échelle est peut-être un peu plus ramassée entre 0,2 et 2 milli-secondes pour le jitter vidéo.

La table ronde a commencé avec un peu de retard, dû à une nouvelle séance de tests pour améliorer l'installation audiovisuelle de *Montréal*. Le déroulement de cet événement a bien mis en avant les difficultés et le temps nécessaire aux utilisateurs pour s'acclimater avec l'environnement coopératif. Les délais induits par le transport du son, même peu importants, contraignent à l'acquisition d'un savoir-faire particulier dans la gestion du droit de parole.

Cette table ronde sur le télé-enseignement rassemblait des enseignants du campus d'*Orsay* et du *CRIM*. L'usage de ces outils étant plus développé au *Québec* qu'en *France*, les enseignants Québécois sont familiarisés avec les délais dans les échanges de paroles. A l'inverse, le manque d'habitude des participants français a provoqué une attente après chaque phrase. Attente correspondant, dans une conversation classique, à la réception d'un acquiescement positif ou négatif, c'est-à-dire à une demande de parole explicite, visuelle ou sonore.

3.2.3. INET'96, Vendredi 28 Juin

La journée de Vendredi prévoyait :

- la séance plénière de clôture de la conférence de l'*IETF*,
- la retransmission et l'enregistrement de l'inauguration mondiale du site *WEB* de la chaîne de télévision Internationale *TV5*.

La retransmission de Jeudi avait montré que la qualité des flux était maintenant stable. Mais qu'une disparité d'affichage apparaissait entre *Orsay* et *Montréal*. Ce jour là, l'image semblait être décodée plus régulièrement, avec une plus grande fluidité, sans perte aussi bien à *Orsay* qu'à *Montréal*. Visuellement, des pertes se caractérisent par l'apparition d'un halo et d'ombres

autour des zones mobiles de l'image. Ce phénomène est provoqué par l'affichage de blocs d'images non consécutifs et codés en différentiel [2]. Notre travail préparatoire a donc consisté à trouver les réglages du débit et du niveau de compression pour chaque type de plan. Nous en avons sélectionné principalement deux :

- le plan général des participants ou éloigné du conférencier, le taux de compression a été fixé à 5, le rafraîchissement maximum à 25 images/s,
- le plan rapproché, le taux de compression a été fixé à 6 et le rafraîchissement maximum à 16 images/s,

Dans les deux types de cadrage, le débit maximum a été fixé à 500 K bits/s. La limitation du nombre de trames/s évite au codeur de boucler rapidement sur des plans fixes et d'envoyer des informations qui ne sont pas réellement utiles, puisque l'image ne bouge pas. Dans la mesure où il n'existe pas, pour le moment, le moyen de distinguer de façon automatique les variations des plans de cadrage, les réglages étaient modifiés à la main par les opérateurs de la régie informatique. Le résultat final, ainsi que les mesures ont validé ce choix technique.

Il existe des expériences d'adaptation automatique du débit des flux aux capacités des traitements des applications clientes. Les résultats de ces algorithmes sont assez décevants sur la qualité des transmissions résultantes. Le principe réside dans l'analyse de la congestion du réseau et du tronçon reliant la source et le client ou les clients distants. Après l'analyse des conditions de réception, l'application génératrice des flux décide d'adapter ses débits aux conditions de congestion. Le but étant de diminuer le débit généré. Dans le meilleur des cas on obtient une oscillation des débits. Plus souvent, les expériences réalisées montrent que le débit du flux ainsi piloté tend à décroître régulièrement. L'image ou le son résultant deviennent inexploitable ou de trop faible qualité. Le deuxième principe, plus philosophique, réside dans le nivellement, par le bas, de la qualité globale de l'événement, conférence, cours, réunion, afin de s'adapter aux plus mauvaises conditions rencontrées. Principe qui ne favorise pas, dans l'état actuel de la technologie, l'activité de coopération humaine. Le débat est à peine ouvert et nous ne le développerons pas dans ce document.

Au cours de cette dernière journée, les pertes ont été limitées. Sans doute en partie, grâce aux derniers réglages liés aux conditions de cadrage réalisées par les cameramans. Ces derniers ont en effet respecté certaines règles pour éviter les brusques variations et également les trop gros plans. Nous ne présentons pas les graphiques de pertes qui ne sont pas d'un grand intérêt par rapport à ceux du 26 Juin.

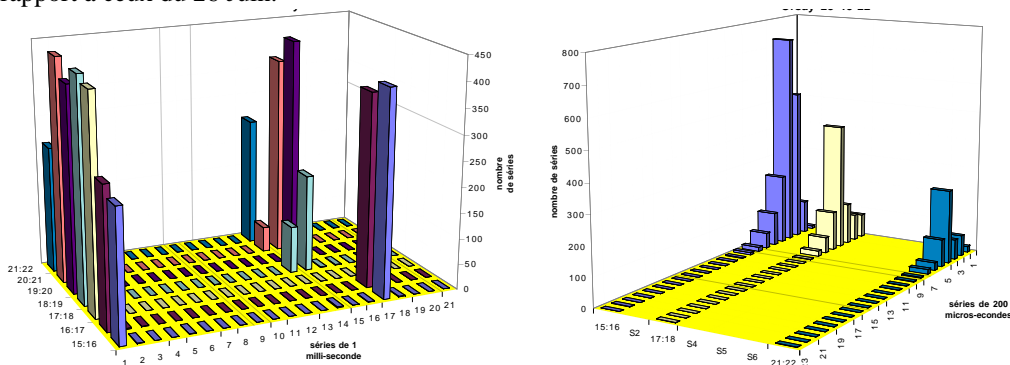


Figure 9 : Orsay 28 Juin, mesures du jitter audio et vidéo

On constate que les dispositions prises ont une influence sur les variations des jitters. Le jitter vidéo est de plus en plus régulier et reste en dessous des deux milli-secondes. Le jitter audio

présente toujours le double pic à 1 milli-seconde puis 17 milli-secondes. Cette dernière valeur évoluant avec l'heure de la journée, pour atteindre 14 milli-secondes.

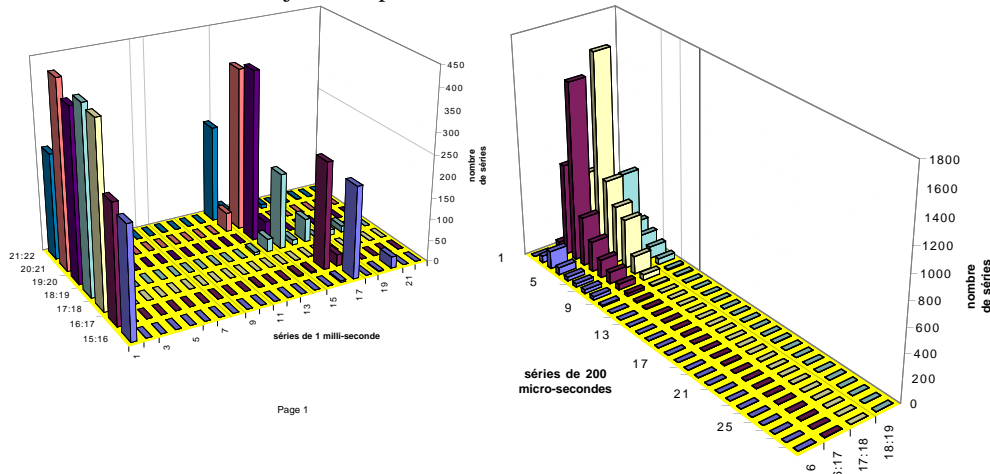


Figure 10 : Montréal 28 Juin, mesures du jitter audio et vidéo

A Montréal, la diminution dans le temps du second pic dans l'audio est encore plus marquée. Il vaut 17 milli-secondes dans la tranche horaire de 15 à 16 heures pour passer à 13 milli-secondes dans celle de 21 à 22 heures. Le jitter vidéo reste régulier.

La retransmission de cette inauguration comportait deux phases. La première s'est déroulée dans un mode proche du séminaire, de Montréal vers Orsay. Il n'y a pas de remarque particulière sur cette partie.

La seconde phase était dédiée aux questions ouvertes de l'assistance présente à Montréal et des personnalités présentes à Orsay. Cette séance interactive a été émaillée d'un phénomène soudain de " boucle audio transocéanique " qu'il a été difficile de maîtriser. Lorsque l'effet Larsen s'est déclenché, la régie audiovisuelle d'Orsay a du réagir immédiatement en baissant les niveaux du retour de la salle de conférence mais sans résultat réel. Il a été nécessaire de couper le décodage du son pour stopper efficacement cette boucle d'écho.

L'explication du phénomène n'est pas aisée, mais cela correspond à une période où l'image et le son ont semblé en synchronisation presque parfaite. Les professionnels de TV5, plus habitués et sensibles à ces détails de qualité subjective ont confirmé cette impression. L'explication la plus plausible est alors celle d'une boucle réactive rendue possible par l'amélioration du délai de transit des paquets. A cette heure avancée de la journée et de la semaine, vendredi 20:30-21:00, les réseaux se désengorgent. Il est probable que la réduction des délais d'acheminement ait permis ce Larsen. Les mesures [figure 9] et [figure 10] montrent effectivement une réduction régulière du jitter avec le temps.

4. Le haut débit pour quoi faire ?

Les expérimentations réalisées par l'équipe TÉLÉSIA ont suivi deux axes principaux :

- l'amélioration de la qualité des flux générés, images animées et son qui conduisent vers une augmentation du débits des flux générés,
- l'adaptation aux conditions rencontrées sur les réseaux " en leur état " qui nous conduisent à contrôler les caractéristiques de ces flux.

Ces deux axes ont permis de réaliser des codeurs-décodeurs vidéo et audio plus performants. Ces briques technologiques sont capables de produire des flux haut débit pour des images et du son de haute qualité, tout en permettant de respecter les limites acceptables sur les réseaux expérimentaux actuels qui sont tous partagés par plusieurs équipes.

Il est en effet possible de tenir compte de la traversée d'un tronçon bas débit par une mise en paquet structurée et une compression plus pertinente. Ces expériences nous ont convaincus, que les technologies développées et mises en œuvre n'avaient de sens que si elles s'inscrivaient dans le cadre d'une activité précise et que la qualité attendue dépend de cette activité. Ainsi, pour le télé-enseignement, l'information passe par le son. Sa qualité est donc un facteur primordial de réussite d'un cours ou du séminaire. Par contre l'image du conférencier est, avant tout, rassurante : "il y a bien quelqu'un qui nous parle". On peut donc la réduire à sa simple expression. Le format *QCIF* avec une fluidité moyenne (10 à 15 images/s) gomme des détails qui, d'après une étude réalisée à *Lyon I* [14] et [15], ne semblent pas être primordiaux.

Durant les tables rondes d'*INET'96*, l'image devait être projetée sur un grand écran pour que les intervenants puissent se sentir intégrés dans cette salle virtuelle. Nous avons vu plus haut que cette contrainte oblige à choisir un format CIF et une plus grande résolution d'image (action sur le taux de compression), tout en conservant, voire en augmentant la fluidité. Cela conduit à une augmentation du débit sur le réseau en terme de volume d'information transmise, de paquets commutés/s et à une réduction souhaitable du délai de transit.

La maîtrise du réseau apparaît donc comme un des éléments clés de ces nouveaux usages des télécommunications. Autant, les problèmes des séminaires retransmis sur *Renater* nous montrent les limites de la technologie de routage, autant la mise en œuvre de l'*IP-ATM* a montré les problèmes d'inter-opérabilité aux frontières : *IP-ATM / IP* routé et du *FMBONE* en tant que réseau logique non intégré. Cette maîtrise est apparue insuffisante pour la liaison *ATM* entre le palais des Congrès de *Montréal* et le centre d'interconnexion de *France Télécom Eurogate*. En effet cette liaison n'a pu être mise en place avec la vision globale du service attendu. Nous n'avons pas disposé d'outils assez fins pour évaluer, en temps réel, les problèmes rencontrés.

Il est manifeste que l'arsenal des outils de mesure reste incomplet pour une évaluation globale. Le jitter est une mesure de variance. Il ne permet pas d'avoir une idée précise des délais effectifs entre la saisie des images et du son en un point et leur restitution à l'autre bout d'une chaîne entièrement informatique.

Nous devons encore faire appel à notre appréciation subjective des délais, pour conclure que les réseaux *IP* classiques routés imposent encore, à ce jour, des contraintes fortes sur une activité ou le dialogue multi-participants prédomine.

A l'heure où l'on imagine déjà de nouvelles applications de vidéo-réunion qui vont permettre de développer un télé-enseignement très interactif, du travail coopératif entre groupe de personnes, des technopoles virtuels et bien d'autres concepts encore, force est de constater que la maîtrise des moyens de communication est insuffisante.

Surtout si l'on cherche à imaginer des outils qui permettraient à l'utilisateur final de fixer les conditions de qualité du service coopératif depuis son interface.

Les nouvelles mises en paquets, que ce soit pour l'audio ou la vidéo, ont induit une diminution de la taille des paquets. Les nouveaux flux véhiculés par l'*INTERNET* font une part plus grande à l'interaction augmentant le nombre de paquets émis par seconde. Il faut s'attendre à une plus grande sensibilité du système de routage face à cette élévation du débit de commutation. Cette évolution donnerait des atouts à l'approche *ATM*.

5. Conclusion

A nos yeux, il reste difficile de réaliser des activités coordonnées à travers un réseau. Nous avons, grâce aux impressions subjectives, une idée de ce qu'il faudrait obtenir pour une plus

grande acceptation, donc diffusion, de cette forme d'activité dans les milieux non dédiés à l'expérimentation de la technologie.

Les premières évaluations permettent de valider certaines approches, tant sur le plan du réseau que celui du système d'information.

Il reste un chemin important à parcourir pour donner à l'utilisateur final les moyens de piloter son environnement de travail coopératif et, à plus forte raison, d'avoir le contrôle global de sa situation.

Cependant, il n'est pas sot de croire et il n'est pas vain d'espérer en la contribution des technologies de réseau haut débit, dans cette nouvelle quête de notre société, toujours à la recherche d'espaces de liberté nouveaux.

6. Remerciements

Nous tenons à remercier particulièrement tout nos partenaires qui ont rendu possible ces expérimentations, qui ont souvent travaillé pour cela hors du cadre de leur fonction :

- le *GIP-RENATER* coordonnateur de l'opération *INET'96*,
- les organisateurs de la conférence *INET'96*,
- Le *CRIM* (Centre de Recherche Informatique de *Montréal*) pour son soutien logistique sur la conférence et la retransmission,
- l'*INRIA* qui assure le support de la vidéo entre postes de travail sur le réseau *IP*, grâce à son logiciel *TÉLÉSIA*,
- *France Télécom*, qui met à disposition les circuits *ATM* français et transatlantiques nécessaires, les opérateurs *Téleglobe*, et *Sprint* (US),
- le *CNET*, qui a mis en place les circuits *ATM* internationaux dans le cadre de *National Host*,
- l'*IUT d'Orsay*, qui organise la logistique d'accueil,
- l'*Association Ile de Science*, qui assure la coordination des événements,
- l'*IDRIS*, qui assure les liaisons spécialisées avec *RENATER* et l'*IUT d'Orsay*,
- l'*Université Paris Sud* opérateur du réseau informatique haut débit du Campus d'*Orsay*,
- l'*Association ARISTOTÉ* qui expérimente et opère les outils de vidéo-conférence,
- l'*École Polytechnique* qui met à notre disposition son réseau et ses moyens audiovisuels,
- les volontaires bénévoles de l'*École SUPELEC* : *Lionel Grenet*, *François Koch*, *Yann Kercadio*, sans qui la réalisation informatique de l'événement *INET'96* n'aurait jamais pu être,
- tous nos partenaires et complices du *FMBONE* et du *MBONE*,
- et tout ceux, qui régulièrement ou occasionnellement, participent aux expérimentations de *TÉLÉSIA*.

7. Bibliographie

- [1] **CANARIE COMMUNIQUÉ**; Vol 2 N°4, Septembre 1994; <http://www.canarie.ca>
- [2] **CCITT; Video Codec for Audio-visual Services at px64kb/s**; CCITT Recommendation H261;
- [3] **David L. Mills; Network Time Protocol, Specifications, Implementation and Analysis**; RFC 1305;
- [4] **Georges Aziza; Denis Joiret; Evaluation du service Transrel-ATM à l'INRIA**; Rapport interne INRIA, Janvier 1995;
- [5] **Grenville J. Armitage, BELLCORE**; Keith M. Adams, U. Melbourne; **How Inefficient is IP over ATM Anyway ?** ; IEEE Network. January/February 1995;
- [6] **H. Schulzrinne GMD Fokus**; S. Casner Precept Software; R. Frederick, Xerox Palo Alto Center; V. Jacobson Lawrence Berkeley National Laboratory; **RTP RFC 1889**;
- [7] **H. Schulzrinne GMD Fokus; RTP Profile for Audio and Video Conferences with Minimal Control**; IETF;
- [8] **Jacques Prévost, Alain Caristan, Sébastien Gibergue; Bilan de l'action de retransmission de la conférence annuelle de l'Internet Society de Montréal à Orsay (Juin 1996)**; Renater Juillet 1996;
- [9] **Jean Yves Babonneau, INRIA**; Alain Caristan, INRIA; **Schéma Directeur des Moyens Informatiques de l'INRIA 1991 - 1994**; INRIA;
- [10] **Jean Yves Babonneau, INRIA**; Gilles Brillet, CNET; Yves Ruggeri, CNET; **National Host, définition et préparation de la conférence de Madrid du 19 et 20 Octobre 1994**; Documents internes INRIA Rocquencourt;
- [11] **Pierre Jardiller; L'organisation humaine du travail**; Collection Que sais je ?; 125
- [12] **Pierre Léonard, INRIA**; Alain Caristan, INRIA; Pierre De La Motte, IFATEC; Andrzej Wozniak, INRIA; **The Tele-seminar Networking with TÉLÉSIA**; <http://magoo.inria.fr/Telesia/telesia-pub.html>; IDC95, Madère;
- [13] **Pierre Léonard, INRIA**; Alain Caristan, INRIA; **Le Multimedia interactifs sur les Inforoutes, expérimentations : TÉLÉSIA, documents de présentation**; <http://magoo.inria.fr/Telesia/DNAC96/>; DNAC96, Décembre 1996, Versailles;
- [14] **Lucia Gradinariu, Marc Jouineau, CISM**; Frédéric Arnaud, Françoise Sandoz GPR-INSA; **Session de Télé-enseignement et bilan pédagogique**; 4^o Journée du réseau ROCAD, 7 Décembre 1995
- [15] **Lucia Gradinariu, Françoise Sandoz, Gérard Beuchot, Patrick Prévot; Interactive Tele-training Using Real-Time Computer Mediated Communication: Usability Evaluation from Experiences**; GRACIMP INSA and Lyon 1 University;
- [16] **R. Atkinson, Naval Research Laboratory; Default IP MTU for use over ATM AAL5: RFC 1626**; IETF, Network Working Group. May 1994;
- [17] **Thierry Turletti, INRIA**; Christian Huitéma, INRIA; **Packetization of H261**; draft-ietf-avt-video-packet-02 September 1994;
- [18] **Ulrich Lang**; Alain Caristan; Lei Wang; Andrea Rozek; **PAGEIN, Pilot Application in a Gigabit European Integrated Network**; R2031/RUS/AWS/DS/L/007/b2 ;
- [19] **Vicky Hardman, UCL UK**; Mark Handley, UCL UK; Martina Angel, UCL UK; Anna Watson, UCL UK; **Reliable Audio for Use over the Internet**; ucl UK;