

La visioconférence : usages et techniques
Par Gilbert Sol, Université Paris 7 - Denis Diderot
Gil.Sol@artemis.jussieu.fr

Dès que l'homme s'est mis à parcourir de nombreuses contrées sur la terre et a connu d'autres semblables, il a voulu communiquer à distance. Ce ne fut pendant plusieurs siècles que par l'intermédiaire de messages écrits à l'exception de quelques cas isolés comme les signaux de fumée chez les Indiens d'Amérique ou le tam-tam en Afrique. On est encore loin de la transmission en temps réel. Il faudra attendre le télégraphe qui n'est d'ailleurs que de la communication par l'écrit. Ce n'est qu'en 1873, que l'invention du téléphone, officiellement par Alexander Graham Bell¹ ouvre la voie aux télécommunications² en autorisant la transmission de la parole en temps réel.

Mais, dans une complète communication à distance, le besoin d'ajouter sa propre image s'est rapidement imposé. Déjà en 1892, moins de 20 ans après l'invention du téléphone, Jules Verne, dans le "Château des Carpates" écrivait : "Deux personnes séparées par des milliers de lieues, causaient entre elles comme si elles eussent été assises en face l'une de l'autre. Elles pouvaient même se voir dans des glaces reliées par des fils grâce à l'invention du téléphone". Cette description nous donne déjà une très bonne définition de la visioconférence ou plutôt de la visiophonie puisqu'elle suppose la transmission du son et de l'image animée sur le réseau de télécommunication de l'époque, le téléphone.

Cependant, il faudra attendre presque un siècle pour concrétiser cette vision, puisque ce n'est qu'en 1972 que le CNET³ établit une première liaison visiophonique sur des liens haut débit entre Paris et Lannion. La première application grand public verra le jour en 1984, lors de l'expérimentation "Fibres optiques" de Biarritz.

La démocratisation et le démarrage de la visioconférence suivra l'installation du RNIS⁴ dont l'ouverture dans les Côtes d'Armor interviendra le 21 décembre 1987 et sa commercialisation sur le territoire français sous l'appellation commerciale "NUMERIS" débutera en 1990. Son extension progressive à presque tous les pays et le respect, par les constructeurs de matériel d'une norme commune, a contribué grandement à l'essor de cette technologie en terme d'usages et de techniques, c'est ce que nous nous attacherons à montrer.

Avant d'aborder les usages, nous ferons le point sur quelques définitions en nous référant aux recommandations de l'UIT-T⁵ qui font figure de normes dans ce domaine. C'est à partir d'exemples pris sur les actions de l'Université Paris 7 - Denis Diderot que nous montrerons comment la visioconférence apporte un soutien efficace en situation d'enseignement ou de Recherche.

Or pour tirer un maximum de profit de cette technologie, le matériel doit être adapté aux fonctions qu'on veut lui faire remplir et intégrer efficacement et en complémentarité les techniques de l'audiovisuel, de codage et de transmission. A cet effet, une présentation rapide

¹ A très peu d'intervalle, Elisha Gray déposa aussi un brevet. Après plusieurs procès, c'est Alexander Graham Bell qui eut gain de cause. Elisha Gray n'en continua pas moins ses travaux et on lui doit en 1888 l'invention du téléautographe pour la transmission de textes manuscrits, invention qui n'eut malheureusement pas de suite.

² Le préfixe "télé" qui induit la distance sous-entend, dans la majorité des cas, l'action en temps réel.

³ CNET : Centre National d'Etude des Télécommunications actuellement renommé France Télécom R & D.

⁴ RNIS : Réseau Numérique avec Intégration des Services dont l'accès de base fournit jusque chez l'utilisateur, sur une ligne téléphonique (2 fils), une liaison numérique avec un débit de 128 kbit/s.

⁵ UIT-T : Union Internationale des Télécommunications, secteur de la normalisation des Télécommunications, organisme de normalisation dans le domaine des télécommunications dont le siège est à Genève.

de la plate-forme de l'Université Paris 7 - Denis Diderot fournira un bon exemple d'intégration des divers médias, des différents types de codage et des réseaux.

Après quelques expériences sur des liaisons spécialisées haut-débit, la visioconférence s'est réellement développée avec la généralisation du RNIS à partir des années 90. Elle a réellement atteint sa maturité sur ce réseau et les derniers codecs⁶ fournissent une qualité convenable mais à des coûts qui ont certes bien diminué, mais qui restent cependant non négligeables puisque à 384 kbit/s (6 canaux B)⁷, la facturation correspond à 6 communications téléphoniques simultanées. Le coût très faible de l'Internet et son internationalisation sont des atouts indéniables, face au réseau RNIS. Cependant, les techniques de transmission sur l'Internet, basées sur le protocole IP (*Internet Protocol*), se prêtent mal à la transmission en temps réel. Dans une dernière partie, après un rappel général sur les normes de visioconférence, nous montrerons comment les lacunes du protocole IP ont été comblées en rajoutant une couche protocolaire dédiée à la transmission du son et d'images animées, en temps réel.

1 - Vocabulaire

L'UIT-T édite dans la série F des recommandations sur la qualité des services à fournir et donne donc les caractéristiques minimales. Nous retiendrons donc quelques éléments qui pourront nous servir de définitions.

Téléconférence (F 710) :

Service de téléconférence: *Service qui fournit les prestations nécessaires à la conduite d'une conférence en temps réel entre des personnes ou des groupes de personnes se trouvant à deux endroits différents ou plus, à l'aide de réseaux de télécommunication.*

Visiophonie (F 720)

Le service visiophonique est un téléservice audiovisuel conversationnel assurant le transfert symétrique dans les deux sens, en temps réel, du son et de l'image animée en couleur entre deux points (de personne à personne) par l'intermédiaire de réseaux. Il faut au moins que dans les conditions normales, la qualité de l'image transmise soit suffisante pour bien reproduire les mouvements fluides d'une personne cadrée «tête et épaules»

Visioconférence (F 730)

Le service de visioconférence est un service de téléconférence audiovisuel en mode dialogue assurant le transfert bidirectionnel, en temps réel, du son et de l'image animée en couleur entre des groupes d'utilisateurs situés en deux ou plusieurs emplacements distincts. Il faut au moins que, dans les conditions normales, la qualité de l'image transmise soit suffisante pour bien reproduire les mouvements fluides de deux ou plusieurs personnes se trouvant dans une situation typique de réunion

La caractéristique commune de ces trois techniques est la transmission en temps réel. Dans le cas de la visiophonie, on rajoute un transfert bidirectionnel, autrement dit, de l'interactivité, et une image animée en couleur. La différence entre la visiophonie et la visioconférence réside dans la qualité de la transmission des mouvements. Très souvent et surtout sur l'Internet, la qualité tient plutôt de la visiophonie que de la visioconférence bien

⁶ Codec : système matériel ou logiciel pour le codage et le décodage des images et du son pour leur retransmission sur un réseau numérique.

⁷ Ces débits sont obtenus en groupant plusieurs accès RNIS.

que les progrès en matière de compression laissent présager une nette amélioration. Cependant, seul le réseau ATM⁸, grâce à son débit élevé et à sa qualité de service est actuellement le seul à assurer une très bonne fluidité et une bonne qualité de l'image.

Nous donnons ci-dessous la liste de l'ensemble des recommandations traitant de visiophonie ou de visioconférence. Les dates indiquées correspondent aux dernières modifications. Le terme de "visioconférence" encore présent en 1992 dans la recommandation G 730 est ensuite remplacé, à partir de 1996 par "conférence multimédia" pour bien montrer qu'on peut rajouter au son et à l'image tout autre type de média.

SERIE F: SERVICES DE TÉLÉCOMMUNICATION NON TÉLÉPHONIQUES

Service audiovisuel

F 700 : Recommandation cadre sur les services audiovisuels multimédias. (11/2000)

F 701 : Directives pour l'identification des spécifications des services multimédias (11/2000)

F 702 : Services de conférence multimédia (07/96)

F 703 : Services conversationnels multimédias (11/2000)

F 710 : Principes généraux applicables aux services de conférence audiographique (03/91)

F 711 : Téléservice de conférence audiographique pour le RNIS (08/93)

F 720 : Considérations générales sur les services visiophoniques (08/92)

F 721 : Téléservice visiophonique pour le RNIS (08/92)

F 723 : Service visiophonique dans le réseau téléphonique public commuté (07/96)

F 730 : Considérations générales sur le service de visioconférence (08/92)

F 731 : Services de conférence multimédia sur le RNIS (07/97)

F 732 : Services de conférence multimédia sur le RNIS large bande (10/96)

F 740 : Services audiovisuels interactifs - Considérations générales (08/93)

Remarque : l'UIT-T déconseille l'utilisation du terme "vidéoconférence" traduction directe de l'anglais "*videoconference*", le préfixe "vidéo" sous-entend une qualité télévisuelle rarement atteinte sauf dans des liaisons ATM.

2 - Usages

La visioconférence intègre dans sa définition, la transmission en temps réel et l'interactivité. Or, les signaux audio et vidéo restent disponibles des utilisations complémentaires comme la diffusion en "streaming video"⁹ et l'enregistrement pour stockage dans une base de données vidéos consultable à la demande. Toutes ces possibilités sont disponibles sur le système de visioconférence de l'Université Paris 7 - Denis Diderot et systématiquement mises en œuvre.

⁸ ATM : (*Asynchronous Transfer Mode*) réseau de Télécommunications large bande à commutation de cellules offrant une très bonne qualité de services en particulier dans le domaine du temps réel.

⁹ "*Streaming Video*" : technique utilisée sur l'Internet pour visualiser une vidéo pratiquement en temps réel. Elle nécessite un serveur qui adapte la qualité au débit du réseau.

Très efficace pour une conversation à distance entre deux personnes, voire un petit nombre de personnes, la visioconférence facilite les échanges pour les préparations de réunions, de cours, de conférences, de colloques, pour la mise en commun de résultats de recherches... Elle évite des déplacements longs et coûteux à un membre de jury de thèse ou à un chercheur devant présenter ses travaux en vue d'un recrutement dans un laboratoire étranger. Elle est aussi très appréciable dans des actes d'enseignement et de recherche comme le montrent quelques exemples d'utilisation, réalisés à l'Université Paris 7 - Denis Diderot.

2.1 - Avec l'ENST (Ecole Nationale Supérieure des Télécommunications) Paris.

Depuis 1999, nos étudiants du DESS "Applications des Réseaux et de la Télématicque" peuvent assister à Jussieu, dans des conditions idéales de qualité audiovisuelle et d'interactivité, grâce à la visioconférence via un lien ATM mis en place lors du projet EFRA, à un cours de 30 heures, le lundi matin de 8h30 à 12h45, de janvier à mars, donné à l'ENST à des élèves ingénieurs de dernière année, sur les réseaux haut débit. Le responsable de ce cours, Daniel Kofman, est un expert des réseaux ATM. Ses compétences l'ont amené à avoir un emploi du temps très chargé et ses interventions à l'Université Paris 7 - Denis Diderot se limitaient à quelques heures, réparties sur une seule journée. Actuellement, nos étudiants bénéficient d'un cours de très haut niveau. Cependant, les travaux dirigés et l'évaluation sont assurés indépendamment par chaque institution.

2.2 - Avec l'Université de Montclair

L'Université de Montclair (*Montclair State University*) dans le New Jersey aux Etats-Unis propose à ses étudiants un module de civilisation française ("*french perspective*"). Depuis 1999, nous organisons une fois par mois, le mercredi après-midi de 15h à 17h, des rencontres par visioconférence sur RNIS en H 320¹⁰, entre les étudiants de l'Université de Montclair et ceux de la maîtrise de Sociologie de l'Université Paris 7 - Denis Diderot. Il s'agit de discussions interactives sur des thèmes choisis par messagerie électronique et préparés par chaque enseignant. Du côté américain, l'apport en apprentissage du français et en connaissance du pays est très appréciable ; la conversation préparée et soutenue avec des Français joue un rôle de catalyseur. Du côté français, la prise de conscience de la mentalité américaine et d'une opinion différente, formulée et discutée en direct sur des sujets d'actualité, enrichit la réflexion sociologique. Devant le succès de cette opération, l'Université de Salt Lake City (Utah) a demandé de participer, pour cette année 2001-2002. Si techniquement, notre système autorise les liaisons multipoints, et donc la possibilité d'assurer des séances communes, nous avons préféré les dissocier car le but est que les étudiants s'expriment. Les liaisons avec l'Université de Salt Lake City s'effectueront donc un autre mercredi du mois.

2.3 - Avec l'Université Claude Bernard de Lyon 1

Pendant deux ans, le professeur Laurence Viennot est intervenu, le jeudi matin, d'octobre à décembre, dans des cours de Didactique de la Physique, assurés simultanément à l'Université Paris 7 - Denis Diderot et à l'Université Claude Bernard - Lyon 1 qui mettait en place un diplôme de troisième cycle dans cette matière. Dans ce cas particulier, la visioconférence a permis de consolider un DEA, de fixer en province un certain nombre d'étudiants et, aussi, de contribuer à la création d'un poste de professeur en Didactique à l'Université Claude Bernard - Lyon 1. Des pourparlers sont en cours pour renouveler

¹⁰ H 320 : norme de visioconférence sur réseau RNIS éditée par l'UIT-T.

l'opération avec l'Université du Littoral - Côte d'Opale, d'ailleurs très bien équipée en matériel de visioconférence.

2.4 - Avec l'ENS (Ecole Normale Supérieure) Lettres de Lyon

L'ENS Lettres a de nombreux contacts avec l'Université Paris 7 - Denis Diderot et l'organisation en commun de cours et de séminaires de troisième cycle, depuis la délocalisation de l'ENS à Lyon, reste possible grâce à la visioconférence. Tous les lundis, de 17h à 19h, à partir du mois de novembre, de 20 à 30 étudiants et chercheurs de Paris, se retrouveront avec une vingtaine de collègues de Lyon, pour des exposés et des discussions. La technique de visioconférence par RNIS en H 320, donne toute satisfaction. Cependant, pour diminuer les coûts de communication, nous envisageons, au cours de cette année universitaire, de tester les liaisons en H 323¹¹ et MBone¹² via l'Internet.

2.5 - Pour l'Agrégation de Lettres Modernes

Dans le concours de l'Agrégation, les étudiants des universités sont en concurrence avec ceux des Ecoles Normales et tout particulièrement dans la préparation de l'oral. Les universités ont parfois des difficultés à maintenir ces préparations. Grâce à la visioconférence multipoint, plusieurs universités dont Paris 3, Montpellier et Paris 7 - Denis Diderot se sont regroupées pour assurer le vendredi matin une préparation à l'oral de l'Agrégation de Lettres Modernes. Des professeurs de chaque université, membres du jury, vont prodiguer leurs conseils et leurs critiques. Les étudiants pourront s'entraîner, présenter leurs leçons d'oral et discuter avec les autres enseignants et collègues sur la qualité de leur prestation. Dans un premier temps, la technique du pont multipoint sur RNIS sera mise en œuvre, mais parallèlement, nous testerons aussi, les liaisons multipoints sur l'Internet.

2.6 - Avec l'Université de Limoges

Nous participons aux enseignements de la licence et de la maîtrise TIC¹³ de la faculté des Sciences de l'Université de Limoges. Les premiers cours sont dispensés en présentiel à Limoges ce qui permet un premier contact. Ensuite, les interventions sont faites depuis Paris par visioconférence. Les étudiants, dont certains sont en "emploi jeune" dans des établissements ruraux équipés de réception satellite, peuvent en complément, suivre les cours du DESS "Applications des Réseaux et de la Télématique" en streaming video ou selon la technique SAT & CLIC par satellite dans le cadre de l'expérimentation du Ministère de l'Education et de la Recherche (les cours sont transmis de Jussieu jusqu'à Vélizy par liaison terrestre à la société Matra-Grollier qui les retransmet vers le satellite HOT-BIRD de TPS ; la réception s'effectue au format IP, grâce à une carte sur un micro-ordinateur).

2.7 - Association "Carrefours Télématiques"

L'association "Carrefours Télématiques"¹⁴ réunit des chercheurs en Sociologie, Communication et Informatique s'intéressant aux impacts, sur la société, des Nouvelles

¹¹ H 323 : norme de visioconférence éditée par l'UIT-T pour les réseaux à commutation de paquets comme l'Internet.

¹² Mbone : (*Multicast Back Bone*) sous réseau de l'Internet permettant grâce à des logiciels spécifiques de faire de la visioconférence entre plusieurs sites simultanément.

¹³ TIC : Technologies de l'Information et de la Communication.

¹⁴ <http://www.artemis.jussieu.fr/ct>

Technologies. Elle compte parmi ses membres des Canadiens, des Marocains et des Algériens. Des séminaires ont lieu tous les mois, le mercredi (pour l'année universitaire 2001-2002) de 18 h à 20h30 ; afin d'atteindre un maximum de membres et de personnes intéressées, ils sont retransmis en temps réel sur l'Internet, aussi bien en streaming video, qu'en mode interactif (MBone et H 323). Leur consultation peut aussi se faire en différé à l'adresse <http://www.artemis.jussieu.fr/ct>.

2.8 - Association "Aristote"

L'association "Aristote"¹⁵ regroupe de grands organismes ou entreprises français intéressés comme acteurs ou comme utilisateurs à l'évolution des télécommunications de transmissions de données. Elle organise en particulier, une fois par mois, le jeudi, un séminaire sur les techniques des télécommunications. Ces séminaires ont lieu à l'Ecole Polytechnique et sont diffusés sur le MBone. Le débit d'accès à l'Internet n'autorise pas d'autres types de diffusion. Nous assurons donc grâce à la plate-forme de VisioP7, la fonction de passerelle. Le séminaire, extrait du MBone, converti en analogique, est renumérisé en temps réel pour être retransmis en "streaming video" sur l'Internet vers les internautes n'ayant pas accès au MBone, et en H 320 sur RNIS vers l'ESMT (Ecole Supérieure Multinationale des Télécommunications) de Dakar.

2.9 - Projet DIM (Diplôme d'Ingénierie Multimédia)

Après plusieurs essais concluants de cours en commun par visioconférence entre l'Université de Valenciennes, l'Université d'Evry, l'INT (Institut National des Télécommunications) et l'Université Paris 7 - Denis Diderot, nous avons élaboré un projet de création d'un diplôme de type DESS, en mettant en commun nos ressources pédagogiques. En effet, nos quatre établissements accueillent des étudiants dans des DESS (DESS "Technologies Nouvelles des Systèmes d'Information" à l'Université de Valenciennes, DESS "Ingénierie Documentaire Multimédia" à l'Université d'Evry, DESS "Applications des Réseaux et de la Télématic" à l'Université Paris 7 - Denis Diderot), comportant dans les programmes certaines similitudes. Le projet consiste donc à réunir nos ressources tant en enseignants qu'en experts de haut niveau afin d'améliorer la qualité de nos formations. Un premier pas a été fait dans ce sens puisque dès le mois d'octobre, tous les mardis après-midi, les étudiants de ces quatre formations suivront le même cours dispensé à tour de rôle par les enseignants de chacun des établissements. Les étudiants pourront ainsi bénéficier des prestations des plus grands spécialistes. Les moyens techniques utilisés seront essentiellement basés sur les techniques de visioconférence sur l'Internet sans toutefois négliger le support du RNIS, en cas de défaillance du réseau informatique.

2.10 Avec l'Université d'Ottawa au Canada

L'Université Paris 7 - Denis Diderot a établi une convention avec l'Université d'Ottawa au Canada dans le domaine des formations à la Communication. Des échanges d'enseignants et d'étudiants ont eu lieu. Quelques difficultés, sont apparues tant du côté enseignant qu'étudiant. En effet, les méthodes d'enseignement et les modes de vie entre les deux pays sont différents. Cependant, ces différences sont aussi un facteur d'enrichissement. Pour conserver ces relations, nous avons fait appel à la visioconférence. Une fois par mois, le mercredi, un cours de Communication est dispensé simultanément dans les deux universités, tantôt par l'enseignant canadien Patrick Brunet, tantôt par l'enseignante de l'Université Paris 7

¹⁵ <http://www.aristote.asso.fr>

- Denis Diderot, Marie-Claude Vettraino-Soulard. Les enseignants et les étudiants peuvent ainsi bénéficier de compétences variées tout en restant dans leur milieu culturel.

2.11 - Avec l'ESMT (Ecole Supérieure Multinationale des Télécommunications) de Dakar et l'UCAD (Université Cheick Anta Diop) de Dakar

L'ESMT (Ecole Supérieure Multinationale des Télécommunications) de Dakar forme des techniciens supérieurs et des ingénieurs en téléinformatique, réseaux et télécommunications. Cette école a été créée et est gérée par plusieurs pays africains : Bénin, Burkina Faso, Guinée, Mali, Mauritanie, Niger, Sénégal et Togo. Nous avons déjà partagé plusieurs colloques et envisageons dès la rentrée, en liaison avec l'Association "Aristote", d'effectuer en commun des cours sur les réseaux et les télécommunications. Le vecteur actuel est la visioconférence sur RNIS en H 320. L'université Cheick Anta Diop désire aussi s'associer à ce projet. L'association Aristote a déjà testé avec succès une liaison à 2 Mbit/s via RENATER (Réseau National de Télécommunications pour la Technologie, l'Enseignement et la Recherche) qui préfigure un lien Internet et pourrait constituer un premier nœud à Dakar pour assurer des visioconférences sur IP.

2.12 - Math en jeans

Une fois par mois, le mercredi de 13h à 17h, depuis deux ans, nous accueillons des lycéens de Camille Sée (Paris 15) avec leur professeur pour des travaux en mathématiques en liaison avec leurs homologues du lycée Doucet de Cluzes (Haute Savoie). Chaque classe prépare un sujet de mathématiques et vient ensuite le présenter par visioconférence. La discussion s'établit très rapidement et les échanges sont très fructueux et motivants. D'ailleurs, ces visioconférences ont établi des liens qui ont permis aux parisiens de passer les vacances de printemps à la neige et au contraire, aux lycéens de Cluze, de venir visiter Paris à Pâques.

2.13 - Dans des conférences ou des manifestations culturelles

En collaboration avec le Ministère de l'Education et de la Recherche et l'association Aristote, nous avons organisé à Jussieu, une journée sur "XML, le format d'échange du XXI^{ème} siècle ; les formats d'échange et standards ouverts : HTML, XML, RDF", le 29 mai 2000. Les exposés étaient présentés dans l'amphithéâtre 24 de l'Université Paris 7 - Denis Diderot, contenant 200 personnes ; une régie mobile assurait le traitement des images issues de 4 caméras. Les images et le son transitaient par la plate-forme multi-réseaux - multitechniques, (décrite en 3) grâce à une liaison ATM haut débit (155 Mbit/s). Les images et le son étaient ensuite redistribués dans la salle 104 où se situait la plate-forme, vers une autre salle de Jussieu via une liaison RNIS Interne en H 320, vers l'ENS de Cachan et l'ENST de Paris en ATM, vers Strasbourg, Rennes, Grenoble et l'Ecole Polytechnique en H 320 via un pont Multipoint RNIS situé à Nancy, sur l'Internet en "streaming video" et sur le Mbone, et via le satellite selon la technique de SAT & CLIC. Cette action nous a permis de définir les éléments à mettre en œuvre pour une régie fixe. Cette



régie intégrant son et vidéo vient d'être concrétisée. L'amphithéâtre 24 (accueillant plus de 200 personnes) est donc, depuis mai 2001, équipé pour retransmettre toute manifestation aussi bien culturelle que scientifique. Grâce à des codecs ATM et à une liaison haut débit en fibre optique, il peut bénéficier de toutes les technologies de compression de la plate-forme. C'est ainsi que, les 3 et 4 mai 2001, nous avons retransmis un spectacle de danse¹⁶ présenté par le groupe chorégraphique "La peau, la pulpe et le noyau" et les étudiants de la filière CMOPC (Conception et Mise en Oeuvre de Projets Culturels) de l'Université Paris 7 - Denis Diderot. La dernière soirée se faisait simultanément sur Paris et Strasbourg sur un principe d'improvisation, par une liaison RNIS en H 320 depuis la plate-forme. Les danseurs de Strasbourg, par écran interposés, évoluaient avec ceux de l'amphithéâtre de Jussieu.

2.14 - Le visio-labo

De nombreux laboratoires ne peuvent recevoir des étudiants, soit pour des mesures de sécurité (manipulation de produits nocifs, chimiques, radioactifs,...), soit à cause de perturbation de l'environnement (salle blanche pour les tests de pollution atmosphérique, bloc opératoire, ...) soit tout simplement, parce qu'il est impossible d'installer un grand nombre d'étudiants devant un écran d'appareil de mesure ou de microscope électronique. D'autre part, certains types d'expériences ne sont réalisés que dans des laboratoires spécialisés, souvent externe au campus universitaire, d'où l'intérêt d'inclure la dimension de transmission à distance.

La solution actuellement mise en œuvre est la réalisation de films sur chaque type d'expérience. Les magnétoscopes se sont généralisés et de nombreuses salles de cours en sont équipées. L'enseignant peut donc s'appuyer sur ces cassettes vidéo pour argumenter son cours. Il est cependant lié à la démarche linéaire de l'auteur et les étudiants ne peuvent intervenir qu'à la fin sous forme d'analyse critique, avec l'aide de l'enseignant. L'expérience ne peut être reprise en concertation avec les étudiants. Elle peut même être déjà dépassée dans sa conception. En effet, l'évolution rapide, dans certains domaines, rend vite désuet les résultats acquis à une époque donnée. Le temps de réalisation et de montage d'un film peut même être, dans certains cas, rédhibitoire. L'avantage de Visio-labo est d'intégrer les deux dimensions : le temps réel et l'interactivité.

Le temps réel supprime tout ce qui peut paraître artificiel conséquence du montage. L'expérience a lieu à l'instant même où elle est montrée. L'étudiant participe à sa préparation, à sa mise en œuvre, à son déroulement et à l'acquisition des résultats, sans fioritures, avec tous les impondérables auxquels se heurte le chercheur. Il pourra assister aux expériences les plus récentes et pertinentes dans tel ou tel domaine. L'étudiant en médecine ou chirurgie assistera à l'intervention rendue délicate par une complication de dernière minute, ... Toutes ces expériences seront d'actualité.

Un autre aspect, et non des moindres, est l'interactivité. Grâce à la visioconférence, l'étudiant, malgré la distance, sera acteur de l'expérience. Il pourra intervenir à tout moment, demander des précisions sur telle ou telle manipulation, remettre en cause certains résultats ; il sera mis en situation de chercheur et non de spectateur.

Plusieurs interventions de ce type ont eu lieu, en particulier avec l'ENS (Ecole Normale Supérieure) de Cachan dans le cadre du projet EFRA¹⁷ pour la retransmission de séances d'utilisation de microscope électronique¹⁸ et lors de la semaine de la science pour

¹⁶ Consultable dans son intégralité à l'adresse : <http://www.artemis.jussieu.fr/danse>

¹⁷ EFRA : Expérimentation Francilienne de Réseau ATM, projet du CNRS soutenu financièrement par le Conseil Régional d'Ile de France en collaboration avec des universités et des centres de Recherche.

¹⁸ <http://www.artemis.jussieu.fr/salon>

une expérience au laboratoire LBHP (Laboratoire de Biorhéologie et d'Hydrodynamique Physico-chimique)¹⁹.

3 - Système de visio-enseignement visio-recherche²⁰ de l'Université Paris 7 - Denis Diderot

3.1 - Historique

La mise en place du système de visio-enseignement visio-recherche commence fin 1996, par l'acquisition d'un meuble de visioconférence SONY TRINICOM 5000 correspondant, à l'époque, à la meilleure adéquation entre coût et usages en environnement de formation (bonne ergonomie, qualité d'affichage, codec à 384 kbit/s, possibilité de recevoir simultanément trois autres sites à 128 kbit/s chacun, sans avoir besoin de faire appel à un pont externe...). Nous lui avons adjoint un ensemble d'auxiliaires pédagogiques : caméra document (banc-titre), micro-ordinateur avec convertisseur professionnel SVGA²¹/PAL²² (pour une bonne qualité du rendu de l'écran du micro-ordinateur), magnétoscope, lecteur de CDI²³, lecteur de vidéodisques, moniteur de contrôle ... le tout regroupé sur un chariot mobile. Ce matériel nous a permis d'effectuer les premiers cours, d'octobre à décembre 1998, à raison de 3 heures par semaine, en didactique de la physique, entre l'Université Paris 7 - Denis Diderot et l'Université Claude Bernard - Lyon 1. Ce système est encore utilisé dans le cadre de visio-réunions ou dans un acte d'enseignement lorsque le professeur exerce depuis Paris à destination d'étudiants distants (en particulier dans le cadre des enseignements de la Maîtrise TIC de l'Université de Limoges).

Fin 1997, nous participons au projet EFRA dans le cadre de la visioconférence sur ATM. Le projet EFRA nous a permis de poser les fondements de notre conception actuelle de la visioconférence puisque nous avons mis en place une régie minimale (sur un seul chariot) intégrant la gestion de la vidéo avec table d'effets (pour, par exemple, afficher deux sources), deux caméras, deux téléviseurs (un pour afficher simultanément



¹⁹ <http://www.artemis.jussieu.fr/fetesience.ram>

²⁰ Mis en place grâce à l'aide financière du Conseil Régional d'Ile de France.

²¹ SVGA : (Super Video Graphics Array) standard d'affichage de la micro-informatique.

²² PAL : (*Phase Alternative Line*) standard européen de télévision couleur à 625 lignes pour 25 images à la seconde.

²³ CDI : (*Compact Disk Interactive*), premier type de disques vidéo remplacés par le DVD (*Digital Versatile Disk*) vidéo.

4 sources grâce à un multiplexeur appelé QUAD, l'autre comme contrôle), une matrice de commutation vidéo (8 x 8), un magnétoscope, deux microphones sans fil, les codecs et le commutateur ATM. Ce dispositif nous permettra d'expérimenter avec l'ENS (Ecole Normale Supérieure) de Cachan le "visio-labo" et d'effectuer les premiers cours en commun avec l'ENST (Ecole Nationale Supérieure des Télécommunications) de Paris à partir de janvier 1999, un module de 30 heures sur les réseaux Haut débit renouvelé depuis, chaque année, et validé à l'Université Paris 7 - Denis Diderot, dans le cadre du DESS "Applications des Réseaux et de la Télématique".

De juin à septembre 1999, nous avons mis en place un serveur de "Streaming video", activé la réception Mbone, testé NetMeeting et conçu la plate-forme multiréseaux-multitechniques qui nous permet de passer d'un réseau à l'autre indépendamment du débit et du type de compression. C'est ainsi qu'en novembre 1999, nous avons transmis sur le Salon de l'Education, Porte de Versailles, une expérience faisant intervenir le microscope électronique de l'ENS de Cachan. Le son et les images sont envoyés en temps réel à notre plate-forme, en ATM du côté de l'ENS de Cachan à 10 Mbit/s en MJPEG²⁴, par NUMERIS à 512 kbit/s en H 320 du côté du Salon. Simultanément, les internautes ont pu assister à l'expérience en RealVideo²⁵ depuis leur micro-ordinateur. Cette démonstration est toujours visible sur l'Internet à l'adresse : <http://www.artemis.jussieu.fr/salon>



Fin 1999, la première version de la plate-forme équipe la salle 104 couloir 46-0. Doublée d'une régie audio-vidéo, architecturée autour d'une nouvelle matrice de commutation vidéo 16 x 16 doublée pour la distribution vers les codecs par l'ancienne matrice 8 x 8 et d'une table de mixage son 16 entrées et 10 sorties, cette plate-forme, capable d'accueillir plus de 8 sites simultanément (indépendamment des techniques et réseaux) assurera le partage et la diffusion des cours ou autres interventions. C'est ainsi que tous les mardis, les cours du DESS "Applications des Réseaux et de la Télématique" seront diffusés en temps réel sur le Mbone, en "streaming video" et à

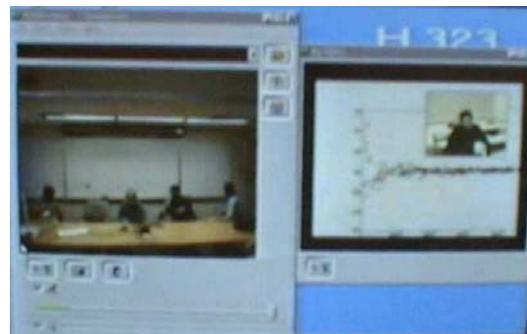


²⁴ MJPEG : *Motion JPEG (Joint Photographic Experts Group)* norme de compression d'images animées considérant chaque image successive comme indépendante ; simple et rapide, elle présente cependant l'inconvénient d'engendrer des débits élevés.

²⁵ RealVideo : technique de "streaming video" mise au point par la société Real Networks.

partir de mars 2000 par satellite dans le cadre de l'expérimentation SAT & CLIC. C'est ainsi que des étudiants de licence et maîtrise TIC de Limoges, en contrat "emploi jeune" dans des établissements scolaires équipés, de l'Académie de Limoges peuvent assister à ces cours. Les cours du DESS sont aussi enregistrés sur un serveur vidéo et les étudiants peuvent les consulter depuis leur poste de travail, lorsqu'ils le désirent pour préparer, en particulier, leurs examens. Tout au long de l'année 2000, les nombreuses expériences réalisées vont nous entraîner à améliorer et à consolider cette plate-forme en particulier dans le domaine de la visioconférence sur l'Internet. Parallèlement, la salle de cours est, elle aussi, adaptée, avec 4 caméras, banc-titre, tableau électronique, écrans supplémentaires pour le professeur et les étudiants... Le nombre de sources traitées simultanément ou susceptible d'être traitées, étant en continuelle augmentation, l'acquisition d'une matrice de commutation vidéo de 32 x 32 s'est avérée indispensable (la matrice 16 x 16 est dévolue à la distribution des codecs en remplacement de la matrice 8 x 8 qui sera utilisée pour l'amphithéâtre) ainsi qu'une deuxième table de mixage son. On peut voir ci-avant l'état de la plate-forme fin 1999 et fin 2000.

Le vendredi 21 avril 2000, nous avons réalisé la première visioconférence de bonne qualité en H 323 sur l'Internet entre notre salle et l'Université de Berkeley. Un étudiant du laboratoire de Physique Nucléaire et Hautes Energies de l'IN2P3 (Institut National de Physique Nucléaire et de Physique des Particules) présentait à son futur jury de chercheurs de l'Université de Berkeley, l'état d'avancement de sa Thèse.



Fin février 2001, nous avons dû déménager la salle de cours et la plate-forme pour l'installer dans la barre 65-66 réhabilitée et dépourvue de toute trace d'amiante. Nous en avons profité pour réorganiser la salle et la plate-forme. Partant du principe que l'enseignant ne doit avoir aucune contrainte (pour que tous acceptent d'être filmés et éventuellement intervenir simultanément sur plusieurs sites) la plate-forme, intégrée à la régie est contiguë à la salle de cours. L'enseignant aperçoit la personne chargée de la régie et peut à tout moment faire appel à lui en cas de besoin. De plus, afin d'augmenter les possibilités de multipoint, nous avons installé un autocommutateur RNIS de 32 accès à 128 kbit/s relié au commutateur général de Jussieu par deux MIC²⁶.



3.2 - Description Technique

Nous avons conçu entièrement notre système de visio-enseignement visio-recherche pour l'adapter au mieux à nos besoins. Son architecture parfaitement modulaire autorise son évolution tant en ce qui concerne l'ajout de nouvelles technologies ou l'adaptation à un besoin spécifique que la capacité d'accueil. Lors de la conception, nous avons toujours gardé à l'esprit nos objectifs : utilisation ergonomique pour des cours dans différentes matières en libérant au maximum l'enseignant et utilisation dans le cadre de formations à la visioconférence ou à l'apprentissage des nouveaux outils de communication.

²⁶ MIC : (Modulation par Impulsion et Codage) liaison sur 4 fils autorisant 30 communications simultanées.

Notre système est architecturé autour d'une régie intégrant la plate-forme qui est l'organe central et opérationnel. Une salle de cours interactive est contiguë, conçue dans un souci didactique, c'est la pièce maîtresse de nos actions. Une deuxième régie, simplifiée, vient d'être installée dans l'amphithéâtre 24 de plus de 200 places. Cet amphithéâtre est relié à la plate-forme par un lien ATM et peut donc ainsi bénéficier de toutes les liaisons et codages disponibles. La salle de visio-réunion équipée avec un codec H 320 sur RNIS, deux micro-ordinateurs avec carte H323/H320 et d'auxiliaires pédagogiques comme la caméra document, un micro-ordinateur portable et un convertisseur PAL/VGA, accède à la plate-forme par le réseau téléphonique numérique interne ou le réseau Internet. Un ensemble mobile constitué d'un codec H320 intégrant caméra et système de saisie du son et d'un codec H 323 connectable à un micro-ordinateur portable par le port USB complète l'ensemble, pour assurer la retransmission d'expériences de laboratoire.

3.2.1 - Plate-forme

La plate-forme a été conçue en fonction de la spécificité des actions de la visioconférence qui fait intervenir plusieurs médias et plusieurs corps de métier (l'audiovisuel, la téléphonie, l'informatique et la pédagogie). Chaque média a été traité séparément. Nous avons réutilisé les chariots qui permettent ainsi de déplacer tout le matériel pour vérifier ou modifier le câblage en fonction des améliorations. La plate-forme dont on peut voir la photo ci-contre, n'est pas encore complètement finalisée et un mobilier mieux adapté sera conçu dès que les aménagements seront stabilisés. Nous allons décrire chaque module.



3.2.1.1 - La vidéo

La plupart des codecs de visioconférence fonctionne au format vidéo PAL (*Phase Alternating Line*) qui est le standard de télévision analogique européen. La qualité est suffisante pour la visioconférence puisque la plupart des codecs dégradent encore l'image dans l'opération de compression. Nous avons donc choisi ce format pour le traitement de l'image.

On peut distinguer sur la



photo ci-contre les divers matériels intervenant dans la fonction de traitement de la vidéo : sous la table d'effet, en bas à gauche, on distingue les deux multiplexeurs d'écrans désignés par "Quad 1" et "Quad 2".

Le "Quad 1" permet la visualisation sur un seul écran, des 4 sources qui seront traitées et sélectionnées par la table d'effets (écran gauche près de la table d'effet). Le "Quad 2" affiche sur un seul écran (écran supérieur gauche) 8 sources vidéos. La sélection d'une de ces sources se fait par appui sur le bouton correspondant du "Quad". C'est pour cette raison que les "quads" ont été placés sous la table d'effets pour plus d'accessibilité. Les autres écrans sont des écrans de contrôle directement reliés à la matrice de commutation vidéo ce qui autorise toutes les combinaisons possibles.

Les deux matrices de commutations ont été installées dans une armoire mais, pour plus de commodités, toutes les connexions ont été ramenées sur l'avant. Nous avons dissocié les fonctions de commutation pour le traitement des sources, qu'elles soient locales ou distantes, fonctions assurées par une matrice de commutation vidéo de 32 entrées et 32 sorties entièrement pilotée par micro-ordinateur, de la fonction de distribution vers les codecs, assurée par une matrice 16 x 16. En effet, nous avons identifié essentiellement trois types de flux dictés par deux situations différentes d'utilisations de la visioconférence. Dans un premier cas, lorsque l'enseignant est à distance, il suffit de lui retransmettre les réactions des étudiants présents sur les divers sites. Un seul flux suffit. Lorsque l'enseignant est dans la salle de cours à Jussieu, il peut utiliser des auxiliaires pédagogiques comme des fichiers sur micro-ordinateur. Les écrans informatiques peuvent être convertis en vidéo au format PAL et mixés grâce à la table d'effets. Cependant, les réseaux actuels, comme l'Internet, et des logiciels applicatifs, permettent le transfert et la manipulation à distance d'images. Deux types de flux peuvent donc simultanément être transportés : l'un qu'on pourrait dénommer "Elève mixé" va transmettre simultanément l'image du professeur et l'image de son écran informatique, pour les sites n'ayant à leur disposition qu'une connexion simple par exemple sur RNIS ; l'autre, identifiée par "Elève simple, sélectionné directement à partir du "Quad 1", ne transmettra qu'une image vidéo, la copie de l'écran informatique étant visualisée via l'Internet.



3.2.1.1.1 - Sources locales

Les principales sources proviennent de caméras. Nous avons choisi des caméras SONY EVI D31 pilotable par logiciel et possédant un zoom de 12. Quatre d'entre elles, équipent la salle de cours, deux orientées vers l'enseignant deux autres vers les étudiants. Ces deux dernières seront asservies aux microphones. L'une d'elle se positionnera automatiquement sur l'étudiant ayant actionné le bouton poussoir du microphone pour poser une question. Ce système est en cours d'installation. Une autre caméra est située dans la



régie, très utile lors d'essais ou lors de présentations ou de démonstrations pour montrer les divers matériels de la régie. La caméra du codec Polyspan est aussi utilisée à cet effet. L'ensemble de ces caméras est actuellement commandé par un logiciel sur micro-ordinateur. Malgré son ergonomie, les déplacements et les modifications de zoom à la souris manquent un



peu de précision. C'est pour cela que nous mettons au point un système de pilotage basé sur un écran tactile pour choisir la caméra et les fonctions associées, et sur deux "Joy-sticks" pour les commandes de positionnement et de zoom.

Trois magnétoscopes (deux VHS et un numérique au format DV CAM) sont disponibles, aussi bien en lecture qu'en enregistrement. Les magnétoscopes VHS savent visualiser aussi bien des cassettes au standard PAL que SECAM. Or, il existe encore des cassettes au format SECAM. Notre régie basée sur le seul format PAL, restitue ces films en noir et blanc. Nous sommes donc obligés d'intégrer un transcodeur PAL/SECAM pour répondre à la demande.

Le professeur a, à sa disposition, une caméra document, couramment dénommé "banc-titre", en remplacement du rétroprojecteur. En effet, le banc-titre autorise la visualisation des transparents. Le professeur peut donc utiliser, sans inconvénient ses anciens supports de cours. Cependant, il réalise rapidement (de lui-même), l'avantage qu'il peut tirer de cet appareil. En effet, il peut montrer des objets opaques, tels que des feuilles simplement sorties de l'imprimante, des extraits de livres et mêmes des objets très petits ou l'écran d'un appareil de test, en tirant profit du zoom très puissant équipant la caméra du banc-titre. De plus, l'image obtenue est directement intégrable à la régie et peut donc être transmise sans difficulté à distance.



L'autre outil, à disposition de l'enseignant, est "l'assistant de conférence digital" qui tient lieu de tableau électronique. Il est constitué d'une barre de capture fixée, par ventouses, sur l'extrémité gauche du tableau blanc.

Elle retransmet les mouvements des feutres à un micro-ordinateur qui retrace en temps réel tout ce qu'écrit l'enseignant. L'écran de l'ordinateur est reconverti en PAL, grâce à un convertisseur, et peut ainsi être traité par la régie comme toute autre source vidéo. La position et la couleur des feutres qui sont insérés dans un étui actif, sont repérées par l'émission d'ultrasons. Un effaceur fonctionne sur le même principe. L'enseignant doit donc assurer sur les feutres une pression régulière pour que le trait retransmis soit bien continu. C'est pour cela que nous avons positionné dans la salle, près du tableau, un téléviseur qui va servir d'écran de contrôle pour le professeur. Les habitudes ne sont donc que très légèrement modifiées.



L'enseignant peut aussi faire appel, pour illustrer ses cours, aux émissions de télévision par satellite, puisque nous avons un abonnement à TPS.

L'enseignant peut aussi faire appel, pour illustrer ses cours, aux émissions de télévision par satellite, puisque nous avons un abonnement à TPS.

3.2.1.1.2 - Sources distantes

Les sources distantes proviennent des différents réseaux via les codecs appropriés (codecs ATM ; codecs RNIS). Les interconnexions via l'Internet sont analysées dans le paragraphe suivant.

3.2.1.1.3 - Sources informatiques

Les sources informatiques proviennent de micro-ordinateurs qui peuvent servir d'auxiliaires pédagogiques locaux ou supporter des codecs matériels ou logiciels transmettant

des images et des sons distants. Dix micro-ordinateurs en régie, se partagent les différentes tâches, sans compter celui qui est réservé au tableau électronique. Comme il était difficile de visualiser tous les écrans, nous avons choisi d'afficher, en permanence, ceux qui étaient destinés à une fonction interactive avec des sites distants et, par un système de diffuseurs et de sélecteurs, les autres sont activés à la demande. Notre souci, a été de pouvoir afficher simultanément deux sources quelconques pour pouvoir les utiliser dans un acte de formation ou de démonstration sur les outils et les usages de la visioconférence. L'ensemble de ces micro-ordinateurs est accessible par seulement deux claviers et deux souris voire même par un seul clavier grâce à un système de sélecteurs. Un deuxième niveau de sélecteurs assure le choix des deux unités centrales dont les écrans doivent être dirigés vers les convertisseurs. Une matrice de commutation VGA assure la rediffusion des sources vers les convertisseurs et/ou vers les deux vidéoprojecteurs de la salle de cours. En effet, un affichage direct en VGA est bien plus net qu'après passage par un convertisseur. Deux entrées de la matrice de commutation VGA correspondent aux sources issues du bureau professeur. Chaque micro-ordinateur a une fonction principale (MBone, H 323, RM). Cependant, tous sont polyvalents et peuvent devenir opérationnels sur MBone, en visioconférence NetMeeting ou comme codeurs de "streaming". Seules les fonctions de pilotage de codecs ATM sont réservées à deux micro-ordinateurs équipés de cartes ATM ainsi que la réception d'Internet par satellite dont la station nous permet de contrôler la qualité de la réception voire de faire une démonstration. Le pont multipoint H 320, hébergé sur une base micro-ordinateur, est aussi commandé à travers ce système.



3.2.1.1.4 - Auxiliaire

Une entrée et une sortie de la matrice de commutation sont notées "auxiliaire". Elles sont facilement accessibles sur le panneau de brassage et sont réservées à l'ajout occasionnel d'appareils soit en test (comme par exemple des codecs MPEG 2 sur IP), soit en complément pour des opérations ponctuelles. Leurs correspondants existent aussi pour le son.

3.2.1.2 - Le son

Le son est la partie la plus délicate dans une visioconférence surtout lorsqu'on utilise des réseaux et des codecs techniquement très différents. L'être humain s'adapte à une image un peu floue ou saccadée alors qu'il se fatiguera très vite si le son est haché, très restreint en fréquence ou variable en volume. De plus, pour éviter les effets de "Larsen" ou d'écho, en multipoint, chaque source doit être envoyée vers les autres mais surtout pas vers elle-même. Nous avons donc prospecté pour trouver une table de mixage répondant à ces critères et autorisant le réglage des volumes aussi



bien en entrée qu'en sortie. Celle que nous avons choisie possède 16 voies en entrée et, en sortie, 4 voies auxiliaires indépendantes, 4 autres voies commutables en deux et une voie de sortie principale. Cette table de mixage équipant la première version de la plate-forme s'est rapidement révélée insuffisante. Nous l'avons donc doublée par une matrice de commutation intégrée sur la même base que la matrice de commutation vidéo 32 x 32 et une autre table de mixage identique, l'achat d'une table de mixage avec beaucoup plus d'entrées et de sorties s'étant avéré très onéreux.

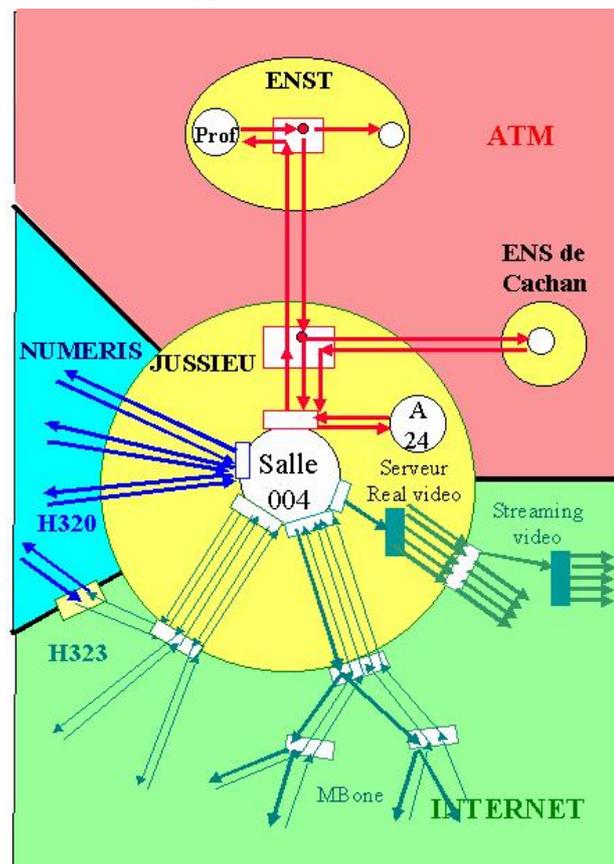
Sur la première table de mixage, on retrouve les sources locales (le son du PC du professeur qui peut vouloir faire une démonstration utilisant le son provenant de l'Internet ou même d'un disque DVD, le son des trois microphones sans fil dont deux cravates, le son issu de la salle de cours), les sources distantes les plus utilisées (liaison ATM provenant de l'amphithéâtre, le Mbone et celles des codecs RNIS Polyspan et Tandberg) ainsi que les sources auxiliaires comme TPS, un magnétoscope ou la sortie d'un des ordinateurs de la régie, sélectionnés via la matrice de commutation. La deuxième table de mixage traite les autres sources distantes.

3.2.1.3 - Les réseaux

Nous avons identifié pour la visioconférence, trois types de réseaux (ATM, RNIS et Internet) et nous avons expérimenté chacun de ces réseaux.

3.2.1.3.1 - L'ATM

Les premières expérimentations ont eu lieu dans le cadre du projet EFRA avec l'ENST et l'ENS de Cachan. Le système est constitué de deux micro-ordinateurs (ATM1 et, ATM2 en secours) gérant l'ensemble du matériel composé de codeurs appelés AVA et de décodeurs nommés ATV. Le codage utilisé est du MJPEG fournissant une excellente qualité avec des temps de compression très brefs à des débits certes élevés, 15 Mbit/s, mais tout à fait compatibles avec les liens actuels (34 Mbit/s avec l'ENST, 155 Mbit/s avec l'amphithéâtre 24). Un des avantages de l'ATM est la diffusion vers plusieurs points sans l'intervention de logiciels ou de matériels supplémentaires. C'est ainsi, qu'un cours effectué à l'ENST peut être retransmis dans une autre salle de l'ENST, vers l'ENS de Cachan, vers l'amphithéâtre 24 et vers la salle de cours 004. La gestion des questions est gérée au niveau de la plate-forme de VisioP7. De plus, l'ensemble des liaisons et le paramétrage de tous les matériels (codeurs et décodeurs) est assurée depuis un seul point. Actuellement, nous possédons 5 codeurs et 5 décodeurs. La liaison vers l'extérieur s'effectue via un codeur et un décodeur. Les autres éléments sont destinés à la liaison avec l'amphithéâtre. En effet, un de nos objectifs est de pouvoir piloter tous les auxiliaires vidéos depuis la plate-forme. De la salle abritant la plate-forme, un codeur transmet à un décodeur



l'image sélectionnée d'un site distant qui s'affichera sur l'écran de l'amphithéâtre. Les sources vidéos de l'amphithéâtre sont au nombre de six (3 caméras, un magnétoscope, un banc-titre et un convertisseur d'écran de micro-ordinateur). Chaque codeur peut recevoir six sources dont la sélection se fait à distance depuis le micro-ordinateur de gestion. Il suffit donc d'installer dans l'amphithéâtre un multiplexeur d'écran (Quad 8) relié à un codeur pour avoir depuis la plate-forme, une vision générale de toutes les sources. Chaque source étant dupliquée sur deux autres codeurs, il sera possible d'en sélectionner deux quelconque pour les traiter avec la table d'effets et retransmettre le résultat sur tous les réseaux disponibles. De plus, les codeurs, équipés d'une liaison série, autorisent le pilotage à distance d'une matrice vidéo ou de tout autre auxiliaire vidéo pour assurer l'affichage sélectif dans l'amphithéâtre. Ce dispositif met à la disposition de toute manifestation scientifique ou culturelle de l'amphithéâtre, la puissance de la plate-forme en limitant de nombre d'intervenants.

3.2.1.3.2 - Le RNIS

Le RNIS (Réseau Numérique avec Intégration des Services) distribué par France Télécom sous le nom de Numéris, est encore très utilisé pour la visioconférence et la norme H 320 est universellement reconnue. C'est pour cela, que la plate-forme fait une part importante à cette technique. En effet, un autocommutateur dédié assure 32 accès relayés vers l'extérieur par deux MIC assurant 60 liaisons à 64 kbit/s. Les fonctions de codage et décodage sont assurées par un codec Polyspan (512 kbit/s et norme H 323 pour les réseaux IP), un codec Tandberg (384 kbit/s), un codec AETA (128 kbit/s) et une carte Vcon sur PC (128 kbit/s et norme H 323). Le tout est complété par un pont H 320 à 5 accès simultanés à 384 kbit/s. La capacité de ce pont correspond à la majorité de nos interventions. En cas de demandes supplémentaires nous faisons appel à un prestataire externe.

3.2.1.3.3 - L'Internet

Au niveau de l'Internet, plusieurs technologies se superposent. Tout d'abord, les terminaux ou logiciels compatibles H 323 : le principe est le même que pour les liaisons sur RNIS en H 320. Essentiellement en point à point, la visioconférence à plusieurs nécessite un pont multipoint. Nous hébergeons actuellement en test, au profit de RENATER, un pont multipoint H 323 à 8 accès simultanés. Une passerelle associée permet le passage du monde RNIS au monde H 323. Sur la plate-forme, les codecs Polyspan et Vcon intègrent la compatibilité H 323. De plus, tous les micro-ordinateurs, grâce au logiciel NetMeeting peuvent émuler un terminal H 323.

La deuxième technologie, idéale pour l'économie de bande passante, est le MBone (*Multicast Back Bone*). C'est celle mise en œuvre dans les milieux universitaires. Le flux audiovisuel n'est dupliqué qu'aux extrémités. Cependant, plus complexe, elle n'est pas disponible partout.

C'est pour cela que nous avons mis en place la diffusion par "streaming video". La capacité de la plate-forme autorise au moins le codage simultané sur trois micro-ordinateurs. En effet, actuellement, trois standards se partagent le marché : RealVideo de la société RealNetworks, Windows Media de Microsoft et Quick Time d'Apple. Le flux codé est transmis au serveur qui peut alors desservir en temps réel les personnes connectées. L'enregistrement peut se faire simultanément. Nous avons choisi de le faire sur la machine assurant le codage. Le transfert est fait sur le serveur après éventuellement un recadrage. C'est ainsi que nous avons mis à disposition des étudiants du DESS "Applications des Réseaux et de la Télématique", tous les cours théoriques du mardi ainsi que les colloques, les séminaires et les interventions dans les laboratoires. Ils peuvent ainsi facilement réviser leurs examens.

3.2.2 - La salle de cours

La salle de cours a été conçue de façon à ne pas modifier les habitudes de l'enseignant. Lors des premiers cours enregistrés, l'enseignant n'était averti qu'à la dernière minute. Tous, sans exception, ont accepté l'expérimentation. L'année suivante, certains professeurs, renseignés par leurs collègues, sont même venus d'eux-mêmes, demander à faire un cours avec enregistrement. Outre les outils traditionnels comme le tableau (amélioré par la barre de capture) et le banc-titre, l'enseignant a, à sa disposition, tous les outils modernes. Nous avons dessiné un bureau intégrant toutes les connexions qui nous ont paru nécessaires. Un micro-

ordinateur avec carte de numérisation vidéo, lecteur de DVD et connexion réseau est disponible en permanence. Des prises électriques, ainsi qu'un HUB USB, un HUB Ethernet 10/100 Mbit/s, et trois connexions écran-souris-clavier autorisent



l'ajout de trois autres micro-ordinateurs, éventuellement portables, avec leurs auxiliaires pour d'éventuelles démonstrations. Un sélecteur quatre ports permet de choisir le ou les micro-ordinateurs qui pourront s'afficher sur l'écran d'ordinateur intégré au bureau et sur les deux autres écrans de projection de la salle, via deux vidéoprojecteurs.

L'affichage peut donc se faire via deux vidéoprojecteurs. L'un est pratiquement opérationnel en permanence. Il autorise la visualisation de l'écran du micro-ordinateur de l'enseignant ou des fichiers transmis à distance. Utilisable en présence de l'enseignant, la forme du bureau interdit à l'enseignant de montrer directement à l'écran avec ses mains, geste

non visible pour les sites distants, lorsque l'on retransmet le contenu des transparents ou de l'écran du micro-ordinateur. C'est aussi pour cette raison que le banc-titre a été placé à l'opposé.

Lorsque l'enseignant se trouve sur un site distant, un écran vient cacher le tableau blanc et la projection de son image se fait face aux étudiants, grâce au deuxième vidéoprojecteur ; le deuxième écran est toujours disponible pour la



visualisation éventuelle d'un fichier image. Les habitudes des étudiants sont donc, elles aussi, peu modifiées. Le système de commutation, aussi bien en mode VGA qu'en mode vidéo autorise pratiquement toutes les combinaisons et peuvent s'adapter à toutes les situations de cours ou de démonstration.

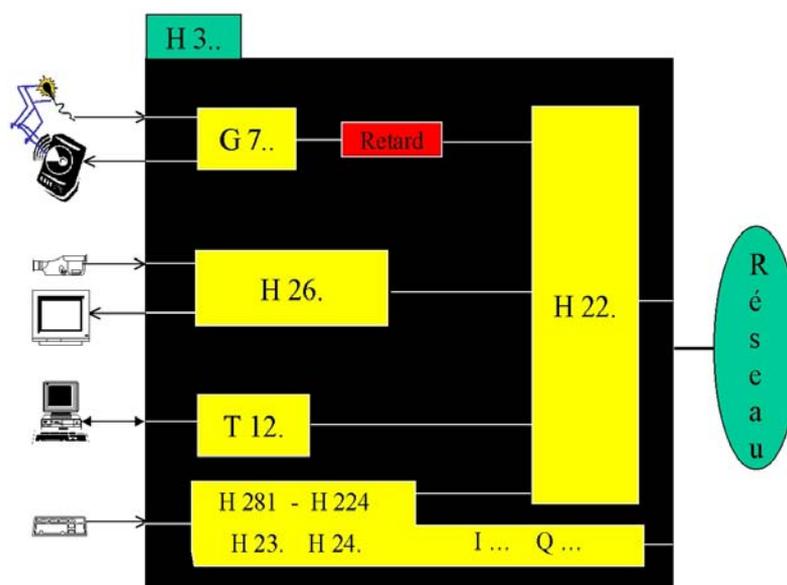
L'enseignant bénéficie de deux écrans de contrôle dont les flux provenant de la régie peuvent être adaptés à la situation. En général l'écran près du tableau, sert à un contrôle de l'écriture alors que celui, au fond de la salle, visualise le ou les sites distants.

Quant au son, un microphone cravate sans fil laisse toute latitude de mouvement à l'enseignant. Deux autres microphones sans fil sont disponibles. Ils sont utilisés par exemple lorsque l'enseignant demande à un étudiant, voire à deux, de venir présenter un rapport ou corriger un exercice au tableau. De leur place, les étudiants désirant poser une question, manipulent un microphone fixe. L'action sur un bouton poussoir, détecté par logiciel, positionne automatiquement l'une des deux caméras. L'apparition en gros plan d'un étudiant sur l'écran du professeur distant présage l'imminence d'une question.

4 - La normalisation

Nous ne donnerons qu'un aperçu rapide de la normalisation. C'est essentiellement l'UIT-T qui s'est chargé du domaine de la visioconférence. Dans une opération de visioconférence interviennent plusieurs médias. La compression de chacun de ces médias est traité dans des recommandation de l'UIT-T spécifiques.

Les codecs se comportent comme des boîtes noires appliquant les recommandations de l'UIT-T. Du point de vue de l'utilisateur, elles comportent des entrées-sorties pour chacun des médias ainsi qu'une prise de connexion au réseau.



Le son est traité dans les recommandations commençant par G 7.. alors que celles correspondant à la numérisation et à la compression vidéo débutent par H 26. La gestion du réseau (début de communication, fin, ...) donné dans les séries I et Q est complétée par des commandes propres à la visioconférence (H 23x ou H 24x. ainsi que H 281 pour le pilotage à distance des caméras dont les codes sont encapsulés dans des trames de type HDLC²⁷ décrites dans H 224 pour la protection contre les erreurs). La transmission de données (transfert de fichiers, annotation de documents à distance, partage d'applications ...) intervient comme complément, en mode optionnel, et est régie par la série T 12x. Les recommandations H 22x décrivent les procédures de multiplexage de ces différents médias. Toutes ces différentes recommandations sont regroupées en fonction des caractéristiques des réseaux et forment les normes génériques H 3xx définissant les codecs. Les tableaux ci-après en donne la liste. Les dates correspondent aux différentes modifications.

²⁷ HDLC : (*High Data Level Link Control*) protocole de transmission chargé de la protection contre les erreurs base des transmissions X 25 de Transpac (réseau français) à commutation de paquets.

Les réseaux

RNIS	H 320	1988 05/99	Systemes et équipements terminaux visiophoniques à bande étroite	64 kbit/s 384 kbit/s 1 920 kbit/s
Réseaux locaux	H 323	11/96	Systemes et équipements visiophoniques pour réseaux locaux offrant une qualité de service non garantie	
		09/99	Systemes de communication multimédia fonctionnant en mode paquets.	
	H 322	03/96	Systemes et équipements terminaux visiophoniques pour réseaux locaux offrant une qualité de service garantie	
Réseau téléphonique	H 324	03/96 02/98	Terminal pour communications multimédias à faible débit	28 800 bit/s 33 600 bit/s
ATM	H 321	03/96 02/98	Adaptation des terminaux visiophoniques H 320 aux environnements RNIS à large bande	64 kbit/s à 1 920 kbit/s (AAL 1)
	H 310	11/96 09/98	Systemes et terminaux de communication audiovisuels à large bande	6 144 kbit/s 9 216 kbit/s n* 64 kbit/s AAL 1 - 5

Le son

H 3 2 0	G 711	1988	300 Hz 3 400 Hz	Modulation par impulsions et codage (MIC) des fréquences vocales	64 kbit/s
	G 728	09/92	300 Hz 3 400 Hz	Codage de la parole à 16 kbit/s en utilisant la prédiction linéaire à faible délai avec excitation par code	16 kbit/s
	G 722	1988	50 Hz 7 000 Hz	Codage audiofréquence à 7 kHz à un débit inférieur ou égal à 64 kbit/s	64 kbit/s 56 kbit/s 48 kbit/s
H 3 2 3	G 723.1	03/96	300 Hz 3 400 Hz	Codeur de signaux vocaux à double débit pour communications multimédias acheminées à 5,3 et à 6,3 kbit/s	5,3 kbit/s 6,3 kbit/s
	G 729	03/96	300 Hz 3 400 Hz	Codage de la parole à 8 kbit/s par prédiction linéaire avec excitation par séquences codées à structure algébrique conjuguée	8 kbit/s
H 3 1 0	ISO/CEI 11 172-3 MPEG1-audio	1993		Codage de l'image et du son associés pour les supports de stockage numérique jusqu'à environ 1,5 Mbit/s. - Partie 3 : audio	
	ISO/CEI 13 818-3 MPEG2-son	1995		Codage générique des images animées et des informations sonores associées - Partie 3 : son	

La vidéo

H 3 2 0	H 261	03/93	Codec vidéo pour services audiovisuels à p x 64 kbit/s	30 im/s 15 im/s 7,5 im/s ...	CIF = FCIF 352x288 (176x144) QCIF 176x144 (88x72) Annexe D image fixe 4CIF 704x576 (352x288)
	H 263	03/96 02/98	Codage vidéo pour communications à faible débit		SQCIF (Sub-QCIF) 128x96 (64x48) QCIF, CIF, 4CIF 16CIF 1408x1152 (704x576)
	H 262 ISO/CEI 13 818-2 MPEG-2	07/95 02/00	Codage générique des images animées et du son associé : données vidéo		MP@ML 720x576 x30i ... HP@HL 1920x1152x60i

Le multiplexage des médias

H 3 2 0	H 221	07/97	Structure de trame pour un canal à débit de 64 à 1920 kbit/s pour les téléservices audiovisuels.
	H 244	07/95	Agrégation synchronisée de canaux multiples à 64 ou 56 kbit/s.
ISO/CEI 13 871 (Bonding)		03/90	Technologies de l'information - Télécommunications et échange d'information entre systèmes - réseaux privés de télécommunications - Agrégation de canal numérique.

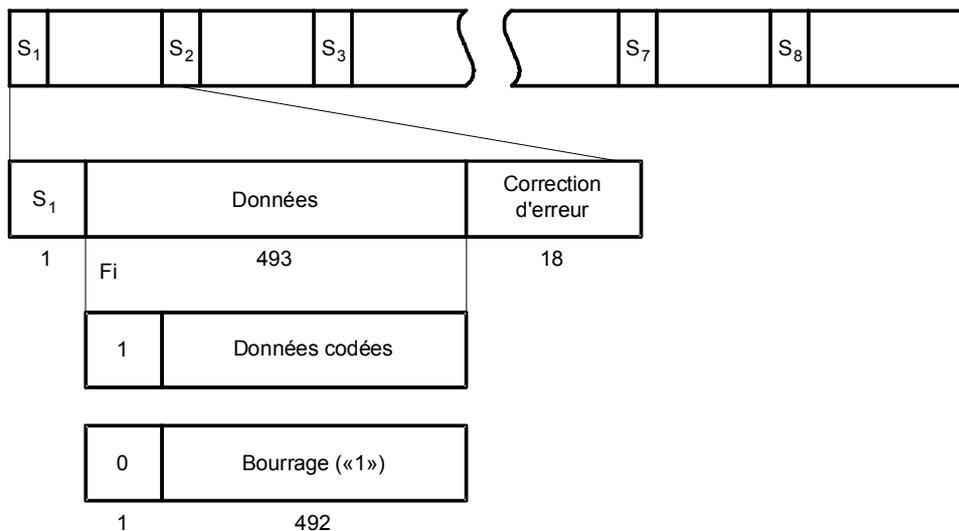
H 3 2 3	H 225-0	11/96	Mise en paquets et synchronisation d'un train multimédia sur réseau local de qualité de service non garantie.
		02/98	Protocoles de signalisation d'appel et mise en paquets d'un train multimédia pour des systèmes de communication multimédia fonctionnant en mode paquets. (RTP, RTCP : RFC 1889 -1890)

H 3 1 0	H 222-0	07/95	Technologies de l'information - codage générique des images et du son associé : systèmes.
	H 222-1	03/96	Multiplexage et synchronisation multimédias des communications audiovisuelles en environnement ATM.

Les modes de transmission des données sur RNIS et sur l'Internet sont fondamentalement différents. Sur RNIS, pour augmenter le débit, il faudra utiliser plusieurs canaux B. On obtiendra alors un débit constant et un temps de propagation fixe et très faible sensiblement proportionnel à la distance. La manière d'agréger les canaux et de multiplexer les médias est fournie par la recommandation H 221, qui n'est pas toujours respectée par les constructeurs de codecs qui souvent préfèrent le *Bonding* (ISO 13 871) plus simple et plus répandue car utilisé en téléinformatique pour des transferts de données, très souvent en secours sur RNIS, de liaisons spécialisées. Par contre, aucun protocole de protection contre les erreurs n'est disponibles dans les canaux B. Ce sera donc aux protocoles des couches supérieures à s'en charger. C'est ainsi que dans H 261 (Compression de la vidéo prévue initialement pour le RNIS), est défini un tramage qui associe la protection contre les erreurs et le remplissage de trames par des bits de bourrage pour assurer la synchronisation, du fait du débit constant.

Dans le cas de la transmission sur l'Internet, du fait de l'utilisation du protocole IP, les données sont mises en paquet et le débit n'étant pas garanti, dans le cas de transmission de données gérées par H 261, le tramage devient inefficace et donc inutile. Les applications de visioconférences supprimeront donc cette étape.

Ordre de transmission → $(S_1 S_2 S_3 S_4 S_5 S_6 S_7 S_8) = (00011011)$



Trame de correction d'erreur

5 - La transmission "temps réel" sur l'Internet

Pour mieux comprendre les contraintes pesant sur la transmission en temps réel à travers l'INTERNET, nous allons d'abord présenter succinctement les protocoles utilisés sur ce réseau.

5.1 - Le protocole de routage IP

Le protocole IP est un protocole sans connexion. Les paquets sont indépendants les

uns des autres et suivent des chemins qui peuvent tous être distincts les uns des autres. Chaque paquet, appelé datagramme, inclut l'adresse de l'émetteur et l'adresse du récepteur. La régulation du flux par les routeurs en est facilitée puisque les datagrammes vont être aiguillés vers les liens les moins chargés. Cependant, la station réceptrice devra pouvoir "stocker" les datagrammes afin de les réordonner puisque le temps de propagation va dépendre du chemin emprunté. Il n'est absolument pas garanti. Le protocole IP, pour cette raison, fournit un service sans connexion. Le service est aussi dit "non fiable". En effet, non seulement le délai, mais aussi la remise d'un datagramme ne sont garantis. En effet, pour diverses raisons, erreurs dans l'en-tête, liaisons saturées, durée de vie écoulee, ..., un routeur peut détruire un datagramme.

De façon plus précise, nous donnons ci-dessous la structure de l'en-tête d'un datagramme IP

0	4	8	16	20	31
Version		Lgr EnT		Type de service	
Identification			Longueur totale		
Durée de vie		Protocole		Drap.	
				Déplacement fragment	
			Total de contrôle en-tête		
Adresse IP source					
Adresse IP destination					
Options IP éventuelles					
Données					

5.2 - TCP (*Transport Control Protocol*) / UDP (*User Datagram Protocol*)

Dans le cadre de la recommandation X25 de l'UIT-T, utilisée par TRANSPAC, la transmission des paquets dans l'ordre et sans erreur est assurée par le réseau. Ce n'est pas le cas pour le protocole IP. Les stations doivent donc assurer ce service. Elles le font selon le protocole **TCP** (*Transport Control Protocol*) correspondant à la couche transport de classe 4 de l'ISO. Ce protocole définit la structure des paquets, appelés **segments**, et les règles de gestion de détection et de retransmission des segments perdus ou en erreur. Il se charge aussi de la remise en ordre des informations pour qu'elles puissent être traitées sans difficulté par les applications telles TELNET ou FTP (*Files Transfert Protocol*).

Cependant, ce type de protocole est mal adapté au temps réel et particulièrement à la transmission de la voix, puisque les demandes de retransmission cassent le rythme de restitution des informations et les rendent rapidement inintelligibles. Le protocole qui lui est donc préféré est **UDP** (*User Datagram Protocol*).

Comme son nom l'indique, UDP est à base de datagrammes comme IP. Il assure ainsi un service non fiable, en mode sans connexion, mais très simple à mettre en oeuvre puisqu'il n'utilise aucun acquittement et n'intervient pas dans le séquençement des données ni même dans un quelconque contrôle de flux. La structure des datagrammes UDP, appelés datagrammes utilisateurs, est réduite au minimum. L'entête n'est constituée que de quatre champs :

0	16	31
Port UDP source		Port UDP destination
Longueur totale du datagramme		Total de contrôle
Données		

Les champs "Ports UDP" source et destination désignent les applications qui sont

concernées par les échanges des messages UDP.

Le total de contrôle est facultatif, mais lorsqu'il existe, il s'applique à tout le datagramme utilisateur englobant les données. Dans ce cas, s'il y a une erreur détectée, il ne délivrera pas ces données à l'application désignée.

5.3 - Protocoles “ Temps réel ”

Le service assuré par UDP se réduit à un routage d'application à application. Il n'est pas suffisant pour assurer un transfert fiable. Il doit donc être suppléé par d'autres couches de protocoles qui assurent en particulier la remise en ordre des informations. Un de ces protocoles est **RTP** (*Real-time Transport Protocol*), repris dans la recommandation H 225-0, spécialement créé pour la transmission de données “ Temps Réel ” comme la voix ou la vidéo.

5.3.1 - RTP (*Real-time Transport Protocol*)

5.3.1.1 - Exemples d'utilisation de RTP

Les exemples qui suivent, décrivent quelques aspects d'utilisation de RTP au-dessus d'UDP ; ils ne sont bien sûr pas limitatifs.

5.3.1.1.1 - Téléphonie et audioconférence

Au niveau d'UDP, deux numéros de port (au-dessus de 5 000) seront spécifiés, un pour la transmission des données sonores via RTP (numéro de port pair), l'autre pour les mécanismes de contrôles engendrés par RTCP (*Real-time Transport Control Protocol*) de numéro impair suivant immédiatement celui de RTP associé. Les données peuvent aussi être chiffrées pour assurer la confidentialité des échanges, mais le cryptage sort de la portée de RTP.

Les données audio sont en général découpées en intervalles de temps de 20 ms (G 723-1 déroge à cette règle puisque les données ne sont compressées qu'après 30 ms), numérisées, codées, puis précédées d'un en-tête RTP et incluses dans un datagramme utilisateur UDP. L'en-tête RTP indique la nature du codage de la source audio (G 711, G 728, G 723-1, ...). Chaque paquet RTP contient cette information. Elle permet en effet la modification dynamique du codage pour s'adapter éventuellement à un nouvel arrivant utilisant un support à bande étroite ou tout simplement, diminuer le débit pour réagir à une indication de congestion du réseau. Ces adaptations sont cependant laissées à l'instigation des applications. RTP assure simplement le transport de ces informations.

Pour permettre à l'application de prendre note de la perte de paquets ou de réorganiser les informations risquant de parvenir dans le désordre, l'en-tête RTP fournit des champs contenant une information de temps (*timing*) et un numéro de séquence afin de maintenir la synchronisation.

Des paquets de contrôle, RTCP, sont émis périodiquement pour connaître à chaque instant le nombre des participants à l'audioconférence, leurs ressources propres et la qualité des liaisons.

5.3.1.1.2 - Visioconférence

Lorsque l'audio et la vidéo sont simultanément utilisés dans la conférence, ils sont transmis dans des sessions séparées de RTP et RTCP utilisant des numéros de port distincts

pour UDP et des adresses de diffusion différentes. Ainsi, certains utilisateurs ne possédant pas les ressources vidéo nécessaires pourront suivre la conférence en audio. Dans le cas d'utilisation des deux médias, RTCP participe à la synchronisation des sources.

5.3.1.1.3 - Mélangeurs (*mixers*) et Traducteurs (*translators*)

Dans les exemples précédents, nous avons supposé que tous les sites recevaient le même format de données. En supposant qu'un certain nombre de participants soient connecté à des réseaux bas débits comme, par exemple, le réseau téléphonique ordinaire, il paraît inconcevable de pénaliser tous ceux qui bénéficient d'un débit important. Aussi, des relais appelés mélangeurs (*mixers*), peuvent être placés aux accès bas débits. Ils récupèrent en entrée, les paquets issus des différentes sources, les resynchronisent et, éventuellement, effectuent un transcodage diminuant le débit, pour reconstituer un seul flux. Le débit est ainsi considérablement diminué. Les indications de provenance des sources seront incluses dans l'en-tête RTP.

Dans le cas où les participants se situent sur des réseaux protégés par des gardes-barrières (*firewalls*), des systèmes, appelés traducteurs (*translators*), sont placés de part et d'autre, et assurent le transcodage des en-têtes IP de façon à franchir le garde-barrière. Des traducteurs peuvent aussi assurer le passage du mode diffusion de groupe au mode simple utilisateur ou encore assurer une adaptation des protocoles de bas niveau pour relier deux réseaux différents.

5.3.1.2 - Format de l'en-tête RTP

0	2	3	4	8	9	16	31
V=2	P	X	CC	M	Type (PT)	Numéro de séquence	
Horodate (<i>Timestamp</i>)							
Source de synchronisation (SSRC)							
Source(s) contribuant(e)s (CSRC)							
...							
Données							

Le premier champ de l'en-tête fournit, sur deux bits, le numéro de **version** de RTP (la version actuelle est la 2).

Le bit suivant **P** (*Padding*), détecte l'éventuelle présence d'octets de bourrage, en fin de paquet, dans le cas où la longueur des données ne correspondrait pas à un multiple de 32 bits. La longueur de cette zone de remplissage est donnée par le dernier octet du paquet lorsque le bit P a pour valeur 1.

Le bit **X** (*eXtension*), donne la possibilité d'étendre le format d'en-tête pour des utilisations particulières ou pour des extensions futures, afin de ne pas avoir à modifier le protocole de base.

Quatre bits formant le champ **CC**, donnent le nombre d'identifiants CSRC contenus dans l'en-tête. Les CSRC (*Contributing source*) correspondent aux sources ayant été "mélangées" lors de la traversée d'un mélangeur.

La signification du bit **M** (*Marker*) dépend du champ suivant **PT** (*Payload type*) définissant le type de données transportées et leur codage (dans le cas de l'audio et de la vidéo, les valeurs de ces champs sont définies dans le RFC 1890). Il sert à marquer des événements significatifs fonctions de la nature des données transportées.

Le **numéro de séquence**, sur 16 bits, va permettre au récepteur de réordonner les

paquets afin de reconstituer le flux initial.

Le paramètre suivant (*Timestamp*) sur 32 bits, donne des indications sur l'instant d'échantillonnage pour la resynchronisation des données. Son incrémentation dépend de la nature de l'application. Elle peut être, par exemple de 1 à chaque paquet, pour le transport du son, lorsque la source et le récepteur utilisent la même fréquence d'échantillonnage.

La **source de synchronisation** (SSRC *Synchronization source*) correspond à la source émettrice, lorsque les données ne sont pas modifiées, sinon c'est l'identifiant du mélangeur qui est fourni. Le traducteur n'intervient pas dans l'identification des sources de synchronisation (SSRC) et des **sources contributantes** (CSRC). Ces deux identifiants sont des nombres choisis de façon aléatoire. La correspondance avec les sources réelles est faite à travers des paquets SDES (Source DEScription) du protocole RTCP.

Lorsqu'un mélangeur retransmet, sur des liaisons de débit inférieur, un flux composé de données issues de sources différentes, il compose l'en-tête du paquet RTP, avec son identifiant en tant que source de synchronisation et indique les identifiants de toutes les sources dont les données ont contribué à la constitution du paquet. Les différentes sources peuvent être, par exemple, deux canaux sons (stéréo) et un canal vidéo.

5.3.1.3 - Extension de l'en-tête RTP

Afin de pouvoir adapter le protocole à de nombreuses applications existantes ou futures, sans avoir à modifier l'en-tête de base, le bit X indique le rajout de champs pour des usages particuliers. Les champs de cette extension faisant suite à l'en-tête de base ont le format suivant :

0	16	31
Défini par le profil, dépend de l'application		Longueur totale de l'extension
Extension d'en-tête		

La longueur de l'extension est donnée en mots de 32 bits. Cette possibilité ouvre la porte à de nombreuses expérimentations sans remettre en cause le standard RTP sur lequel pourront s'appuyer les nouvelles applications.

5.3.2 - Protocole RTCP (*Real-time Transport Control Protocol*)

Comme nous l'avons vu précédemment, RTCP, lorsqu'il est utilisé sur UDP, utilise un port différent de RTC. Les deux fonctions de transport de données et de contrôle de transmission sont donc distinctes. RTCP prend en charge les fonctions qui ne sont pas assurées par les protocoles sous-jacents IP et UDP, comme par exemple, la surveillance de la qualité de la transmission, la gestion de flux ou la gestion des participants dans des applications de travail en groupe.

5.3.2.1 - En-tête des paquets RTCP

Le premier mot de 32 bits de chaque paquet RTCP ressemble à celui des paquets RTP. Il a la forme suivante :

0	2	3	8	16	31
V=2	P	RC ou SC	Type (PT)	Longueur totale	

On retrouve le numéro de **version** de RTP et le bit **P** (*Padding*) indiquant l'éventuel rajout de bits de bourrage pour obtenir un multiple de 32 bits. Le troisième champ de 5 bits, précise le nombre de blocs d'information contenu dans le paquet ; selon le type de paquet, il se nomme **RC** (*Report Count*) ou **SC** (*Source Count*). Le **type de paquet PT** (*Packet Type*) est défini sur 8 bits. Les 16 derniers bits donnent la **longueur totale** en mots de 32 bits.

5.3.2.2 - Paquets SDES (*Source Description*)

L'association entre un identifiant SSRC ou CSRC et la description complète de la source est effectuée par un paquet **SDES** (*Source Description*) de type 202. Il va permettre en particulier à l'application d'afficher les caractéristiques des sources (nom, email, numéro de téléphone, ...), d'associer plusieurs médias provenant d'une même source et de tenir à jour les noms des participants à une même conférence. Un paquet SDES est décomposé en blocs décrivant chaque source à l'aide de caractéristiques normalisées identifiées par un champ de 8 bits suivi de la longueur de la description sur 8 bits. La liste se termine par au moins un octet nul (éventuellement complété par des bits à 0 pour que l'ensemble de la description d'une source occupe un nombre d'octets multiple de 32). Les types de caractéristiques actuellement définis sont :

- **CNAME** (*Canonical name*) de code 1, toujours présent, permet d'associer les différents identifiants SSRC ou CSRC à une source réelle définie généralement par un nom et une adresse de domaine. Alors que le SSRC ou CSRC d'une source peut être modifié puisqu'il est aléatoire, en raison d'un conflit ou de la réinitialisation d'une application, le CNAME identifie de façon unique le participant.
- **NAME** (*user name*) de code 2, est le nom qu'utilisera l'application pour l'affichage des participants. Il identifie en général l'utilisateur contrairement au CNAME qui correspond plus précisément à la station.
- **EMAIL** (*Electronic mail adresse*) de code 3.
- **PHONE** (*Phone number*) de code 4.
- **LOC** (*Geographic user location*) de code 5, dépend de l'application ; il peut se limiter à une ville ou donner l'adresse complète.
- **TOOL** (*Application or tool name*) de code 6, identifie le nom et la version de l'application générant le flux de données.
- **NOTE** (*Notice*) de code 7, sert à transmettre de courts messages comme le titre d'une conférence ou le status d'une source.
- **PRIV** (*Private extension*) de code 8, réservé à des applications particulières ou expérimentales.

D'autres noms et codes pour les caractéristiques pourront être attribués en fonction des futurs développements.

Un exemple de format de paquet SDES est donné ci-après.

0	2	3	8	16	31
V=2	P	SC	Type = SDES = 202	Longueur totale	
SSRC 1 ou CSRC 1					
CNAME=1		Longueur		...	
...		NAME = 2		Longueur ...	
...					

...		
...		NUL
SSRC 2 ou CSRC 2		
CNAME=1	Longueur	...
...		
...		
...		NUL
SSRC 3 ou CSRC 3		
...		

5.3.2.3 - Surveillance de la qualité de la transmission.

La surveillance de la qualité de transmission des données s'effectue à l'aide de paquets RTCP appelés **SR** (*Sender Report*) ou **RR** (*Receiver report*). Le format de ces deux paquets ne diffère, outre le numéro de type, que par l'inclusion d'un bloc de 20 octets correspondant aux caractéristiques de l'émetteur en ce qui concerne le paquet SR.

5.3.2.3.1 - Format d'un paquet RR (*Receiver report*)

0	2	3	8	16	31
V=2	P	RC	Type = RR = 201	Longueur totale	
SSRC de l'émetteur du paquet					
SSRC 1					
Fraction perdue		Nombre total de paquets perdus			
Numéro de séquence étendu du dernier paquet RTP reçu de SSRC 1					
Gigue interrarrivée					
Horodate du dernier rapport SR (<i>LSR Last SR</i>)					
Délai depuis le dernier rapport SR (<i>DLSR Delay since last SR</i>)					
SSRC 2					
...					
...					
...					
Extension éventuelle					
...					

Un paquet RR, après l'en-tête RTCP et l'identifiant de l'émetteur, est structuré en bloc fournissant un rapport de réception pour les différentes sources détectées. Chaque bloc indique:

- le taux de perte des paquets RTP (nombre de paquets perdus / nombre de paquets prévus), pour la source considérée, depuis l'expédition du dernier rapport,
- le nombre total des paquets perdus depuis le début de l'émission,

- le numéro de séquence du dernier paquet RTP reçu, sur 16 bits, suivi du nombre de cycles depuis le début de l'émission,
- une estimation de l'écart moyen de la différence d'espacement de deux paquets au niveau du récepteur par rapport à l'émetteur,
- l'instant de réception du dernier rapport SR concernant cette source (en temps absolu approché sur 32 bits au format NTP²⁸),
- Le délai depuis la réception du dernier rapport SR.

Le paquet peut se terminer par un bloc d'extension généré pour une utilisation particulière.

L'ensemble de ces indications contribue à maintenir une bonne connaissance de l'état de la liaison pour permettre aux applications concernées d'adapter leur débit ou le type de codage des données multimédias.

5.3.2.3.2 - Format d'un paquet SR (*Sender Report*)

Le format d'un paquet SR correspond à celui d'un paquet RR avec insertion d'un bloc d'information sur l'émetteur.

0	2	3	8	16	31
V=2	P	RC	Type = SR = 200	Longueur totale	
SSRC de l'émetteur du paquet					
Horodate (partie entière du format NTP)					
Horodate (partie décimale du format NTP)					
Horodate conforme à l'unité des paquets RTP					
Nombre total de paquets RTP émis depuis le début de la transmission					
Nombre d'octets utiles émis dans les paquets RTP depuis le début de la transmission					
Identique aux paquets RR					
...					
...					

Les champs relatifs à l'émetteur, fournissant des précisions de temps, vont permettre au récepteur d'obtenir des indications précises sur les instants d'émission des paquets RTP, tout en ayant sa propre horloge. La correspondance peut ainsi être établie entre l'unité de temps transporté par les paquets RTP et le temps réel, en particulier elle permettra, par exemple, de calculer la fréquence d'échantillonnage du son dans le cas d'une application d'audioconférence.

5.3.2.4 - Autres paquets RTCP

D'autres types de paquets sont définis par RTCP :

- BYE** de code 203 précisant les identifiants des sources (SSRC ou CSRC) quittant l'application,
- APP** de code 204 pour des extensions non encore prévues dans ce protocole.

²⁸ NTP (*Network Time Protocol*) : la date est donnée en secondes à partir de 0 heure UTC, le 1^{er} janvier 1900, sur 64 bits, 32 bits pour la partie entière et 32 pour la partie décimale. Dans les protocoles RTP et RTCP seuls les 32 bits médians sont utilisés, c'est-à-dire 16 pour la partie entière et 16 pour la partie décimale.

Conclusion

Jusqu'à présent la visioconférence a surtout emprunté, comme réseau de télécommunications, le RNIS qui offre une bonne qualité pour la transmission en "temps réel". L'Internet, basé sur le protocole IP est au contraire très sensible à ce type de données et en particulier dans le domaine de la voix. Les protocoles RTP et RTCP assurent certes, la synchronisation mais ne sont pas en mesure d'assurer la garantie de transmission des paquets dans des délais compatibles avec le "temps réel". L'augmentation des débits n'est pas non plus entièrement satisfaisant ; il faut agir directement sur les routeurs pour qu'ils traitent en priorité les datagrammes transportant des données sensibles au "temps réel". C'est ce que nous expérimentons avec RENATER, dans le cadre du projet DIM présenté dans une autre conférence de JRES 2001.