

Rapport INRIA 1994 — Programme 1
Réseaux à haut débit, Réseaux ouverts

PROJET RODEO

3 mai 1995

PROJET RODEO

Réseaux à haut débit, Réseaux ouverts

Localisation : *Sophia-Antipolis*

Mots-clés : ALF (1), appel de procédure distante (1), ASN.1 (1), contrôle de transmission (1), ILP (1), ISO-OSI (1), protocole de communication (1), réseau à haut débit (1), réseau informatique (1), RNIS (1), vidéoconférence (1).

1 Composition de l'équipe

Responsable scientifique

Christian Huitema, Directeur de recherche, Inria.

Responsable permanent

Walid Dabbous, Chargé de recherche 1, Inria

Secrétaire

Gisèle Zimol

Personnel Inria

Jean-Chrysostome Bolot, Chargé de recherche 2, Inria

Christophe Diot, Chargé de recherche 2, Inria

Ingénieur

Jean-Patrick Giacometti, Ingénieur Inria à partir du 1^{er}
Octobre 1994

Chercheurs post-doctorant

Torsten Braun, boursier H.C.M., depuis le 1^{er} Septembre 1994

Ellen Siegel, boursière ERCIM du 1^{er} Avril jusqu'au 30 Septembre 1994

Chercheurs doctorants

Thierry Turetti, boursier Inria

Philipp Hoschka, boursier H.C.M.

Isabelle Chrisment, boursière Inria

Claude Castelluccia, boursier Inria

Jean Noël Pic, boursier DRET, depuis le 1^{er} Octobre 1994

Andrés Végas Garcia, boursier du gouvernement mexicain

Stagiaires

Kamal Okba, stagiaire ENIM, du 15 mars au 15 Septembre 1994

Aline Russeil, stagiaire DESS/ESSI, du 15 Avril au 15 Octobre 1994

Hugues Devries, stagiaire EURECOM, du 10 Janvier au 31 Juin 1994

Anna Hoglander, stagiaire ENSIMAG, du 1^{er} Juin au 30 Septembre 1994

Hugues Crépin, stagiaire DEA informatique, du 1^{er} Novembre au 31 Juin 1994

Sandro Mazziotta, stagiaire de recherche, du 1^{er} Juin au 30 Septembre 1994

Autres personnels

Anthony Richards, collaborateur extérieur de UTS à Sydney en Australie à compter du 20 Novembre 1994

2 Présentation du projet

Le but du projet RODEO est de définir et d'expérimenter les mécanismes qui permettront de tirer parti des réseaux à 1Gbit/s et plus.

Pour réaliser cet objectif, nos thèmes de recherche s'articulent autour de deux axes:

- Le contrôle de transmission à haut débit.
Nous étudions et mettons au point des protocoles de contrôle de transmission à haute performance qui intègrent la gestion des services multipoint, le partage des ressources, le contrôle d'erreur, etc. Ces protocoles sont en particulier mis en oeuvre dans notre application de vidéoconférence IVS (INRIA Videoconference System).

- Le développement d'applications à haut débit.
Des travaux précédents ont conduit à la mise au point du compilateur ASN.1 MAVROS, et au développement d'algorithmes rapides pour la gestion de la présentation des données pour transmission entre machines hétérogènes. Nous poursuivons ce travail par la prise en compte de spécifications formelles plus détaillées: en enrichissant les connaissances disponibles par le compilateur, on autorise de nouvelles optimisations.

Ces deux axes sont appelés à converger dans le cadre de la nouvelle architecture "Application Layer Framing". La prise en compte de spécifications détaillées devra permettre de générer sur mesure des modules gérant aussi bien présentation et synchronisation que le contrôle de transmission.

Ces travaux se poursuivent en collaboration avec des projets français, européens, et internationaux. Les travaux sur une architecture de protocoles haute performance se poursuivent avec des partenaires tels que University College London, SICS (Suède), et UTS (Australie). Nos travaux sur l'application vidéoconférence ont donné lieu à plusieurs actions de coopération internationales, notamment avec le projet européen MICE (avec UCL, GMD, SICS, et l'Université d'Oslo). Nous participons au projet Value PASSWORD sur la sécurité dans les réseaux. Nous collaborons également à des actions de standardisation des protocoles de réseau et des applications multimédia dans le cadre de l'IETF (Internet Engineering Task Force). On note que le projet adresse les problèmes de communication indépendamment des supports de transmission: réseaux locaux, fibres optiques, ATM, etc.

3 Actions de recherche

3.1 Contrôle de transmission

Les mécanismes de contrôle de transmission limitent le taux d'utilisation des ressources d'un réseau (liaisons de communication, CPU et mémoire dans les routeurs, etc) en contrôlant le flux des paquets dans un réseau. Ils ont pour but l'utilisation optimale et équitable des ces ressources par l'ensemble des utilisateurs du réseau.

Les travaux cette année ont poursuivi des axes de recherche commencés les années précédentes. Les efforts ont porté particulièrement sur le développement de mécanismes de contrôle pour applications multimédia (audio-vidéoconférence, tableau blanc partagé) sur Internet. 1) l'analyse et la modélisation de mécanismes de contrôle, 2) le développement de mécanismes de contrôle de transmission pour applications audio et vidéoconférence sur Internet, 3) le développement de mécanismes de contrôle d'erreur pour applications de type tableau blanc partagé, et 4) le développement de mécanismes permettant d'offrir des garanties de qualité de service.

3.1.1 Analyse de mécanismes de contrôle

Participants : Jean-Chrysostome Bolot

Les mécanismes de contrôle de transmission utilisés dans les réseaux actuels sont des mécanismes à fenêtre qui limitent pour chaque source (ou utilisateur) le nombre de paquets en transit dans le réseau. Par exemple le mécanisme utilisé dans le réseau Internet est le mécanisme à fenêtre de TCP, qui modifie la taille de la fenêtre en fonction de la charge dans le réseau (estimée par le taux de pertes des paquets). Nous poursuivons des travaux effectués dans le cadre d'une collaboration CNET-INRIA (projet MISTRAL) sur le développement de modèles analytiques pour évaluer la performance de TCP sur le réseau RENATER. Ces modèles doivent prendre en compte l'aspect dynamique de TCP, i.e. le fait que la taille de la fenêtre varie en fonction des changements d'état du réseau. Les modèles stochastiques classiques conduisent à des analyses compliquées. Nous avons donc développé des modèles simples en temps continu (modèles fluides).

En parallèle avec ces travaux de modélisation, nous avons continué nos mesures sur le réseau Internet. Nous observons en particulier les délais et pertes de paquets dans le réseau. L'accent cette année a été mis sur la compréhension du processus de perte de paquets envoyés à intervalle régulier dans le réseau. Ce problème est important car il permet d'évaluer la qualité d'un flux audio envoyé sur Internet (puisque les paquets audio sont envoyés en général à intervalle régulier). Nous avons observé, et l'analyse d'un modèle analytique (le modèle $D+D^X/D/1/K$) a confirmé, que la plupart des pertes sont isolées tant que le taux de charge du réseau est faible ou modéré. Les pertes deviennent groupées seulement à charge élevée. Ces résultats nous ont conduit en particulier à proposer l'utilisation pour la transmission de la voix sur Internet de méthodes de correction d'erreur par anticipation. Une première implémentation dans le logiciel de vidéoconférence IVS (voir ci-dessous) a permis de vérifier que de telles méthodes sont susceptibles de compenser des taux de pertes élevés. Le travail consiste maintenant à développer et implémenter des méthodes adaptatives, c'est à dire dans lesquelles le mécanisme de correction d'erreur s'adapte aux conditions (i.e. au processus de perte en cours) à un instant donné dans le réseau. Ces travaux sur les mesures et l'analyse de trafic vont continuer pendant les 3 prochaines années dans le cadre d'un contrat INRIA-CNET.

3.1.2 Contrôle de transmission pour applications multimedia sur Internet

Participants : Jean-Chrysostome Bolot, Hugues Crépin, Hugues Devries, Christian Huitema, Thierry Turletti

Nous avons réalisé un système de vidéoconférence à bas débit sur le réseau Internet qui utilise le protocole de transport UDP (User Datagram Protocol) et l'extension IP multicast définie par le RFC 1054. Cette application appelée "INRIA Videoconference System" (IVS) permet d'effectuer des audio/vidéo conférences avec plusieurs participants via leurs stations de travail. Elle permet aussi de retransmettre des séminaires et conférences via le MBONE (réseau multicast expérimental au dessus d'IP). Les applications de type vidéoconférence nécessitant beaucoup de bande passante, les données audio/vidéo sont comprimées avant leur émission.

Nous avons choisi d'implémenter un logiciel de codage/décodage vidéo qui suit la norme H.261 du CCITT (codec vidéo pour services audiovisuels à $p * 64kb/s$). Plusieurs techniques sont employées conjointement pour augmenter la performance globale du système: une boucle de prédiction appliquée dans la dimension temporelle, une transformation en cosinus suivie d'une quantification des coefficients transformés, et d'un codage entropique (codage de Huffman).

Plusieurs formats d'encodage sont possibles: le format QCIF (176 * 144 pixels), le format CIF (352 * 288 pixels) et le format SCIF (704 * 576 pixels). Le logiciel a été porté sur les plate-formes suivantes:

- Sun SPARCstation w/ SunOS 4.1.x et SunOs 5.3 avec les cartes vidéo d'acquisition VideoPix, Parallax et SunVideo.
- DECstation 5000/25 w/ ULTRIX V4.2 avec la carte vidéo VI-DEOTX.
- Silicon Graphics Indigo w/ IRIX 4.0.5 avec les cartes vidéo IndigoVideo, Galileo vidéo ou Vino vidéo.
- HP9000/7xx w/ HP-UX 8.0x avec la carte vidéo VideoLive.

Une version pour PC w/ FreeBSD 2.0 est en cours de réalisation.

La norme H.261 a été à l'origine conçue pour une utilisation sur des réseaux à commutation de circuits comme le RNIS. Pour être utilisée sur des réseaux à commutation de paquets comme l'Internet, une adaptation a donc été nécessaire. Un schéma de paquetisation du flot vidéo H.261 a été soumis comme Draft Internet au Working Group "Audio-Video-Transport" de l'IETF. Il a été adapté dans le cadre du projet MICE par les codecs hardware Bitfield et GPT du commerce, ce qui nous permet une interopérabilité avec ces derniers.

La réalisation d'un codec logiciel a été rendue possible grâce à l'augmentation de la puissance des machines. Un codec réalisé en logiciel présente de nombreux avantages par rapport à un codec H.261 réalisé en hardware. Le codec hardware a un coût plus élevé et ses caractéristiques sont figées, tandis qu'un codec logiciel peut être facilement adapté à de nouvelles machines plus puissantes, et il est de plus facilement contrôlable. Ce dernier point est important car on désire en effet contrôler le rythme d'émission du codeur en fonction des capacités de transmission du réseau. Un tel mécanisme de contrôle nécessite un algorithme de détection de capacité disponible intégré au transport des données, et un

algorithme d'asservissement du codeur adapté à la transmission d'information multimédia. Le réseau Internet actuel ne propose aucun service permettant d'obtenir la capacité disponible ou le taux de charge du réseau. IVS inclut donc un mécanisme permettant d'estimer la capacité disponible dans le réseau et ce, quelquesoit le nombre de participants dans la conférence. Cette estimation permet d'ajuster le débit de sortie du codeur vidéo de manière à satisfaire le maximum de participants dans la conférence.

Pour cela le codeur dispose de plusieurs paramètres, à savoir :

- le rythme de rafraîchissement d'image,
- le quantificateur: Un quantificateur plus grossier comprimera davantage les données au dépend de la qualité de l'image,
- le seuil de détection de mouvement. En modifiant ce seuil, on change la sensibilité de détection de mouvement et donc le nombre de pixels à encoder d'une image à l'autre,
- la précision de la transformée en cosinus. On peut réduire le nombre de coefficients dans le domaine transformée.

Mais cette adaptation pose des problèmes délicats en raison de l'hétérogénéité inhérentes des récepteurs. D'autres techniques peuvent être utilisées pour éviter que certaines branches de l'arbre de distribution multipoint restent congestionnées.

Les résultats obtenus sur l'Internet indiquent que l'adaptation du codeur en fonction des conditions du réseau permet d'offrir une application de vidéoconférence de qualité raisonnable sans avoir à utiliser les mécanismes de contrôle d'admission et/ou de réservation de ressources qui sont généralement considérés comme nécessaires pour ce genre d'applications. Cependant nous ne pouvons pas *garantir* une qualité donnée à l'utilisateur. Les travaux sur ce thème sont décrits dans la section 3.1.4.

D'autres travaux cette année ont visé à fournir des mécanismes qui permettent l'adaptation de débit pour le codeur audio d'IVS. Malheureusement, il est beaucoup plus difficile de contrôler de façon continue le débit d'émission d'un codeur audio que d'un codeur vidéo. La solution retenue a été de fournir un ensemble de codeurs audio, et de choisir à un instant donné le codeur approprié (c'est à dire le codeur qui génère la meilleure qualité de voix possible pour une bande passante maximum donnée) en fonction de la capacité (i.e. de la bande passante) disponible dans le réseau à cet instant. Les codeurs disponibles actuellement

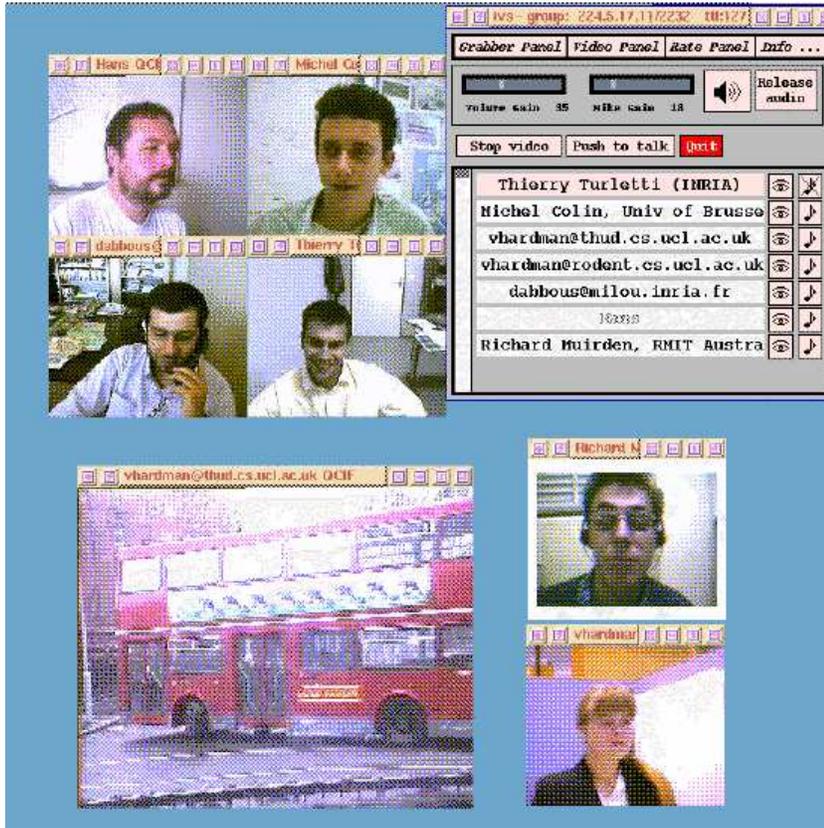


Figure 1 : Démonstration IVS entre partenaires du projet MICE

sont des codeurs PCM (64 kb/s), ADPCM (16, 32, et 48 kb/s), et LPC (5 kb/s). Il reste cependant à effectuer l'adaptation de débit de façon automatique comme cela est fait pour la vidéo.

La transmission de la voix sur Internet pose d'autres problèmes spécifiques. En particulier, les paquets de voix sont très sensibles aux pertes (qui engendrent des blancs dans la parole) ainsi qu'aux délais (qui diminuent l'interactivité) mais aussi à la variation des délais (appelée gigue, et qui engendre une parole hachée). Nous avons donc développé des mécanismes qui permettent de compenser les effets néfastes de ces problèmes. En ce qui concerne les pertes, nous avons développé des mécanismes de reconstruction de la voix basés sur des techniques de correction d'erreur par anticipation (ces travaux ont été décrits dans la section 3.1). En ce

qui concerne les délais, nous avons développé un mécanisme de compensation de gigue qui permet en quelque sorte d'absorber à la destination les variations de délai des paquets de voix.

Nos travaux sur IVS sont effectués en coopération avec des partenaires européens dans le cadre du projet MICE. Il s'agit d'une coopération avec des partenaires universitaires et industriels, notamment UCL, GMD, et SICS, afin de définir un standard de vidéoconférence pour les réseaux de recherche.

Les sources et binaires du logiciel IVS sont disponible par ftp anonyme a partir de `zenon.inria.fr` dans le répertoire `"/rodeo/ivs/last_version"`. Une description du logiciel est disponible sur le serveur Web de l'INRIA via l'URL:

`<http://www.inria.fr/rodeo/personnel/Thierry.Turletti/ivs.html>`

3.1.3 Tableau blanc et transmission multipoint fiable

Participants : Walid Dabbous, Kamal Okba

Le développement de protocoles de transport multipoint fiables est une brique nécessaire pour la programmation des systèmes distribués et des bases de données distribuées. Cependant, cela est une tâche compliquée car les choix de conception sont beaucoup plus délicats à effectuer que dans le cas d'un protocole point à point. En effet, hormis le problème de gestion de groupe qui pourrait être traité par l'application, trois difficultés majeures restent à résoudre en ce qui concerne le transfert de données: garantir la fiabilité de la transmission, remettre les données dans l'ordre à la réception et contrôler l'envoi des données sur le réseau (le contrôle de transmission).

Beaucoup de protocoles de transport multipoint ont été conçus. Ces protocoles assurent d'une part la fiabilité et d'autre part l'ordre (causal ou total) correspondant à une sémantique spécifique de la diffusion. Pourtant, les premières implémentations de ces protocoles avaient des problèmes de performance et surtout d'échelle ce qui a entretenu l'idée que la transmission multipoint fiable était intrinsèquement non efficace. En effet, l'absence d'un support "réseau" pour la transmission multipoint avait imposé son implémentation en une séquence d'envois de paquets en point à point. L'avènement d'un service de routage et d'acheminement multipoint sur l'Internet a permis l'adaptation de certains protocoles afin de profiter de cette possibilité. La transmission des don-

nées en multipoint couplée avec l'utilisation d'acquittements négatifs ou (Nacks pour Negative ACKnowledgemntS) résout le problème de la fiabilité de façon élégante. La remise dans l'ordre des données est assurée par une numérotation des messages et un mécanisme d'horloge globale.

Ceci n'est pourtant pas suffisant pour concevoir un protocole de transport multipoint efficace. En effet, la politique d'envoi des Nacks et des retransmissions est un facteur déterminant des performances du protocole. Si certains choix statiques permettent le fonctionnement du protocole dans le cadre d'un réseau local, il est très difficile de gérer la transmission sur un réseau maillé comme l'Internet sans l'ajout de mécanismes permettant l'adaptation de la transmission à la topologie et aux autres caractéristiques du réseau (délais, taux de pertes). Nous avons développé un protocole de transport multipoint fiable appelé de façon générique RMP (ou Reliable Multicast Protocol). RMP est basé sur un système d'acquittements négatifs. Chaque message ou paquet de données émis est caractérisé par un numéro de séquence, une estampille de temps et l'identification de l'émetteur. Dans le but d'avoir des performances acceptables ainsi qu'une bonne tolérance aux pannes aucun mécanisme de jeton explicite ne sera utilisé. Une station qui a des données à transmettre au groupe envoie un paquet en multipoint au groupe. Quand il n'y a pas de pertes, un paquet émis ne nécessite aucun paquet additionnel. La détection de la perte d'un paquet implique une génération d'un Nack, en vue de demander la retransmission du paquet. Le choix du moment d'envoi du Nack ainsi que le contrôle des réponses au Nack sont déterminants pour les performances du protocole. En effet, un Nack signifie une congestion éventuelle dans le réseau, et la réponse au Nack ne doit pas être la retransmission du paquet par un grand nombre de stations, sinon ça aggraverait la congestion actuelle. Il faudrait pouvoir adapter le rythme d'envoi de paquets de sorte à permettre une communication fiable avec des délais de transmission aussi faibles que possible. C'est justement sur ce point que nous avons focalisé nos recherches en 1994: le choix des mécanismes de contrôle de transmission du protocole permettant d'obtenir un fonctionnement efficace et performant du protocole. En effet, s'il est vrai qu'en l'absence des pertes l'utilisation des Nacks permet d'améliorer les performances surtout en multipoint, il y a un risque d'implosion de Nacks dans le cas où plusieurs stations détectent la perte d'un paquet. Ce problème est aggravé si les stations émettent leur Nack immédiatement après la détection de la perte. Cette "synchronisation" engendrera des rafales de

Nacks vers la station source. Les performances de RMP dépendent donc de la méthode de gestion des Nacks utilisée: Quand est ce qu'un Nack sera envoyé? Qui répondra aux Nacks et aux autres messages de contrôle (comme les requêtes de joindre le groupe)?

Deux mécanismes de contrôle ont été étudiés:

- la définition d'une hiérarchie dans le groupe permettant de répartir les stations afin d'éviter des retransmissions multiples par les stations recevant le même Nack. Cette hiérarchie est basée sur une variable locale à chaque station appelée "profondeur" désignant l'ordre de la station dans une file d'attente distribuée. Cette variable est incrémentée lors de la réception des messages de données et mise à zéro lors de l'envoi d'un message par la station. Sa mise à jour ne nécessite donc pas l'envoi de messages de contrôle spécifiques.
- l'écoute du réseau avant d'émettre couplée avec une randomisation des réactions des stations aux messages de contrôle reçus. Cette technique à la CSMA permet de favoriser la réponse par une station proche de la station ayant envoyé le Nack.

Cette étude a montré les limites de ces deux mécanismes. Le mécanisme de contrôle par profondeur impose dans certains cas une correction des erreurs par une station éloignée du fait que cette station avait émis en dernier. D'autre part, l'approche à la CSMA ne permet pas d'éviter les retransmissions multiples ou "collisions" dans certains cas.

Nous étudions actuellement un mécanisme de contrôle permettant de combiner les avantages des deux approches décrites ci-dessus: en plus de l'écoute du réseau et de la randomisation de l'envoi des messages de contrôle, la profondeur est utilisée afin d'éviter les collisions. Une modélisation ainsi que des expérimentations nous permettront d'évaluer les performances de cette approche.

Ce protocole est testé actuellement avec une application de tableau blanc partagé (appelé Mscrawl). Nous allons aussi étudier la possibilité de configurer le protocole afin d'avoir un contrôle de transmission optimisé sur mesure pour d'autres applications.

3.1.4 Contrôle et garanties de qualité de service

Participants: Jean-Chrysostome Bolot, Andres Vega

Une évolution importante dans le monde des réseaux est l'apparition d'applications intégrant différents types de données qui génèrent des trafics avec des contraintes différentes, la priorité étant donnée dans certains cas à l'intégrité des données (transfert de fichiers), dans d'autres cas à la régularité des délais de transit (transfert de voix numérisée).

L'intégration de différents types de trafic a souvent été mise en avant pour justifier l'utilisation de mécanismes de contrôle d'admission et de réservation de ressources. Nos travaux sur IVS (voir ci-dessus) ont montré qu'il est en fait possible d'intégrer ces trafics sur l'Internet tout en obtenant une qualité de service raisonnable sans ces mécanismes, mais au contraire en adaptant le débit des différents types de trafic en fonction des conditions du réseau (Ces travaux sont décrits dans la section 3.1.2).

Cependant, l'adaptation de débit seule ne suffit pas à garantir une qualité de service minimum aux applications. Il faut pour cela mettre en œuvre des mécanismes d'allocation et/ou de réservation de ressources qui garantissent qu'une application ait à sa disposition une connexion avec une bande passante minimum ou un délai maximum. Les travaux cette année ont porté sur les mécanismes d'allocation de ressources. Le mécanisme retenu est en fait une discipline de service, c'est à dire une méthode qui détermine à chaque instant dans un routeur à quel instant et dans quel ordre les paquets en attente dans ce routeur seront transmis. La discipline en question est appelée "weighted fair queueing" (WFQ), ou "attente équitable pondérée". Ce mécanisme permet en fait de séparer le trafic arrivant dans un routeur en plusieurs flux. A chaque flux est associé une fraction minimum des ressources disponibles. La séparation des flux peut se faire de différentes façons, par exemple sur la base des machines source et destination, des types d'applications, etc.

Ces travaux sont mis en oeuvre et évalués sur un réseau de test. Pour l'instant, ce réseau de test est constitué de PCs Unix (utilisant la version domaine public FreeBSD de ce système d'exploitation) reliés entre eux par des liaisons série et Ethernet. La mise en oeuvre de ce réseau n'a pas été sans problème, du fait principalement de problèmes causés par le manque de drivers Ethernet pour le logiciel FreeBSD. Cela nous a conduit à développer un tel driver, qui utilise des techniques efficaces comme la prédiction d'interruption et le parallélisme pour obtenir de très bonnes performances. Ce driver a d'ailleurs été choisi pour faire partie de la nouvelle version de FreeBSD (version 2.0) qui devrait être disponible à la fin de l'année.

La discipline de service WFQ a été implémentée dans le réseau de test. L'implémentation a été optimisée afin de réduire au maximum le nombre d'opérations nécessaires au traitement d'un paquet. Les travaux en cours visent à évaluer précisément en quelle mesure l'utilisation de cette nouvelle discipline de service permet de diminuer les taux de pertes de paquets et les délais subis par les paquets dans les routeurs. Les premières mesures indiquent, comme prévu, qu'elle diminue effectivement pertes et délais, et qu'elle offre une bonne isolation entre flux. Ce qui permet, par exemple, qu'une transmission audio ou vidéo ne soit pas perturbée par des surcharges de trafic temporaires créées par des applications sans contrainte temps réel comme les transferts de fichiers. Il reste cependant à effectuer une comparaison détaillée entre la discipline WFQ et la discipline utilisée actuellement qui est la discipline FIFO (premier arrivé premier servi). L'impact de la discipline sur la qualité de service offerte aux applications qui utilisent le réseau sera mise en évidence sur notre réseau de test grâce à l'utilisation des applications multimédia développées dans le projet, en particulier l'application de vidéoconférence IVS.

3.2 Codage de Présentation à haut débit

Participants : Christian Huitema, Philipp Hoschka, Sandro Mazziotta

La transmission de données sur un réseau nécessite souvent une conversion du format des données appelée codage de présentation (ou "marshalling"). Actuellement, le codage de présentation est reconnu comme étant une des opérations les plus coûteuses dans un pile de protocole. Les routines de marshalling sont générées par les compilateurs de talon ("stub") à partir des spécifications d'interface d'une application. Les compilateurs actuels génèrent des routines qui nécessitent typiquement plusieurs milliers d'instructions de processeur par message. A titre de comparaison, les implémentations des protocoles de transport optimisés ont besoin seulement d'une centaine d'instructions par message.

Les performances des routines de marshalling sont influencées par deux facteurs majeurs:

- La stratégie de génération du code utilisée par le compilateur de talon ("stub"). Les deux stratégies normalement utilisées sont la génération d'un code intreprété ou la génération d'un code compilé.

- Le format de données utilisé comme format d'échange. Les formats les plus connus aujourd'hui sont: BER (Basic Encoding Rules) utilisé dans SNMP et tous les standards OSI, NDR (Network Data Representation) utilisé dans OSF/DCE, XDR (eXternal Data Representation) utilisé par Sun et PER (Packed Encoding Rules) proposé comme simplification du BER.

Un problème classique pour la génération de talons est de trouver un compromis entre la taille du code généré et les performances des routines de conversion:

- La génération de tables qui seront interprétées au moment de l'exécution résulte en un code qui a une taille compacte, mais aussi une exécution lente (approche interprétée),
- La génération directe du code des routines de conversion résulte en un code qui a une exécution plus rapide, mais aussi une taille du code élevée (approche compilée).

Pour générer du code qui est suffisamment performant et ayant un taille compacte, il faut combiner ces deux approches. L'approche compilée sera appliquée seulement pour les types d'une spécification d'interface qui seront utilisés fréquemment en cours de l'exécution. Cette approche est efficace si la distribution de l'usage de types en cours d'exécution présente des propriétés de localité. Nous avons fait des expériences qui confirment cette hypothèse.

Pour automatiser le compromis entre l'approche interprétée et compilée, il faudrait disposer d'informations sur la fréquence d'utilisation dynamique des types, d'un modèle pour prédire les performances et d'un algorithme pour décider si un certain type devrait être compilé ou interprété.

Nous avons fait des expériences qui indiquent qu'une heuristique simple et souvent très exacte existe pour prédire les types les plus souvent utilisés au moment de l'exécution. Cette heuristique est basée sur l'analyse de l'arbre de types d'une spécification d'interface pour prédire le flot de contrôle dans les routines de marshalling. Pour prédire les performances, un simple modèle de prédiction était développé. Ce modèle utilise le nombre de composants d'un type donné multiplié par la fréquence d'utilisation dynamique de ce type comme indicateur du temps et de la taille du code pour ce type. Avec les informations pour chaque type et une limite sur la taille totale du code de marshalling, on peut

formuler le problème de la génération du code de marshalling comme un problème Knapsack. L'algorithme de décision sur la stratégie optimale peut donc être un des algorithmes et heuristiques déjà existants pour résoudre les problèmes Knapsack.

Le deuxième paramètre ayant une influence sur les performances des routines de marshalling est le format des données utilisé sur le réseau. La plupart des travaux précédents sur l'optimisation des routines de marshalling essaient de trouver un format d'échange qui est proche du format de données local sur une machine. Cette approche a des problèmes de compatibilité évidents. Donc, nous avons décidé de suivre une approche complémentaire, qui consiste à analyser de façon plus détaillée la cause du problème de performance des formats existants, et de proposer des algorithmes qui peuvent améliorer les performances du marshalling pour ces formats existants. Cette analyse doit prendre en compte les architectures des systèmes haute performance d'aujourd'hui comme l'utilisation de processeur RISC et la grande différence entre la vitesse du processeur et la vitesse de la mémoire.

Une première constatation émanant de ces travaux est que la génération du code de marshalling devrait être basée sur la séquence de mots qui contient le message plutôt que sur la séquence des champs du message. Suivant cette approche "orientée-mot", nous étions capables de générer un code de marshalling pour le format BER avec une performance proche du format XDR (200 Mbit/s pour BER sur un Sparc 10 Modèle 40 pour un séquence de 32 entiers correspondant à un en-tête TCP/IP. Les routines "orientées-champs" générées par Mavros ont permis d'obtenir des performances de 35 Mbit/s). Le débit étant défini comme étant le rapport entre la taille du code généré et le temps de codage. Un problème reste à résoudre: comment contrôler l'explosion de la taille du code avec l'algorithme de génération de code utilisé dans ces expérimentations. Des études sur ce point sont en cours, en collaboration avec l'Université d'Arizona, Tucson.

Pour agrandir notre base d'expérimentation, nous avons étendu le compilateur ASN.1 MAVROS par un générateur de talons pour les formats de données PER ("Packed Encoding Rules") et XDR ("eXternal Data Representation"). En plus, nous avons étendu le compilateur `c2mvr` pour générer des routines de marshalling directement à partir de "fichiers headers" C, sans passer par une représentation intermédiaire en ASN.1. Ce

compilateur sera mis dans le domaine public. Une première version est prévue pour début 1995.

Une page W3 avec des informations sur le langage ASN.1 était développée: (<http://www.inria.fr/rodeo/personnel/hoschka/hoschka.html>). Plus de 500 accès dans les 4 premières semaines de sa mise en place indiquent que cette page est devenue très vite un moyen d'information d'Internet standard pour la communauté d'utilisateurs d'ASN.1.

3.3 Architecture flexible de protocole haute performance

Participants: Walid Dabbous, Christophe Diot, Christian Huitema

Avec l'apparition des réseaux à haut débit, la bande passante de la couche d'accès au milieu ne représente plus un goulot d'étranglement. Ce sont plutôt les mécanismes de traitement de données au niveau transport et présentation qui nécessitent une charge de calcul importante au niveau des processeurs. D'autre part, des applications multimédia distribuées sont en cours d'introduction. Ces applications ont des besoins variés qui ne correspondent pas nécessairement aux services offerts par les protocoles de communication standard ou du moins pas à la qualité de service fournie par ces protocoles. En même temps, les performances des stations de travail ont augmenté avec l'apparition des architectures à jeu d'instructions réduit (RISC) mais pas à la même cadence que les bandes passantes des réseaux dans les dernières années. L'accès mémoire est devenu coûteux par rapport au cache et aux registres, et cette différence devrait s'accroître dans les années à venir. Outre les fonctions de contrôle, le traitement d'un protocole renferme aussi des fonctions de manipulation de données où les données sont lues en mémoire, traitées ou "manipulées" et sauvegardées de nouveau en mémoire éventuellement (e.g. codage de présentation, calcul du total de contrôle au niveau transport, chiffrement, compression). Nos propres observations, confirmées par des mesures de performances publiées par ailleurs, montrent qu'il est possible d'obtenir des gains de performance en intégrant les opérations de manipulation de données afin de minimiser les accès mémoires. Cette technique appelée ILP (Integrated Layer Processing ou Traitement en Couche Intégrée) devrait être complétée par un mécanisme visant plus de flexibilité: l'application devrait être impliquée dans le contrôle de transmission en effectuant une "Mise en Unités de Données Applica-

tion” (Application Level Framing ou ALF) permettant d’appliquer tous les mécanismes de manipulation sur ces trames de façon intégrée i.e. en une seule fois si possible. La mise en œuvre de ces concepts pour des applications différentes nécessite qu’on puisse générer pour chaque application une implémentation intégrée des mécanismes de protocoles fournissant le service requis par l’application. Afin d’automatiser ce processus, il est nécessaire d’avoir un langage de spécification qui permettra de décrire les besoins de l’application.

Notre approche s’appuie sur la spécification des besoins des applications en ESTEREL et sur un outil appelé “compilateur de protocoles” basé sur les concepts ALF/ILP et qui va générer automatiquement le protocole de communication adapté à l’application à partir de sa spécification.

Nous avons suivi le plan suivant:

- Validation du choix du langage ESTEREL pour la spécification des besoins des applications.
- Expérimentation de l’applicabilité du principe ALF à certaines applications.
- Etude et expérimentation de la technique d’implémentation ILP.
- Conception et développement du compilateur de protocoles.

3.3.1 Spécification de protocoles avec ESTEREL

Participants : Claude Castelluccia, Isabelle Chrisment, Christophe Diot, Ellen Siegel

Un protocole de communication est un système réactif; il peut être vu comme une “boîte noire” activée par des événements en entrée (paquets entrants) et produisant des événements en sortie (paquets sortants). Il est donc tout à fait concevable de spécifier un protocole de communication en ESTEREL. Outre les avantages liés à l’utilisation d’un langage formel, les intérêts d’une telle démarche sont les suivants :

- Le compilateur ESTEREL construit (entre autres) des automates qui décrivent le module compilé. Représenter un protocole de communication par un automate est tout à fait naturel. Grâce à la notion de relations, le compilateur ESTEREL permet d’optimiser les automates qu’il produit. Il est donc possible de construire un protocole en fonction des besoins de l’application.

- ESTEREL est un langage réactif. Il ne possède pas d'horloge interne. Ce sont les événements externes qui déclenchent l'exécution des modules ESTEREL ; cette exécution étant instantanée. En d'autres termes, ESTEREL préserve l'atomicité dans le traitement d'un événement; ce qui signifie que lorsqu'un nouvel événement arrive, on est certain d'avoir terminé le traitement de l'événement précédent. Charge au compilateur de générer un automate fini déterministe tout en autorisant la spécification d'actions parallèles.

Plusieurs expérimentations ont été effectuées afin de valider le choix du langage ESTEREL. Nous avons étudié les différentes approches de spécification d'un protocole de communication en ESTEREL. Nous avons en particulier spécifié un protocole de contrôle d'erreur. Cette expérience nous a permis de déterminer les conditions d'utilisation d'ESTEREL pour la description de protocoles de communication, ainsi que les limites de ce langage.

Nous avons aussi spécifié et implémenté un protocole de transport proche du protocole TCP. La spécification obtenue est très modulaire, chaque fonctionnalité du protocole est spécifiée par un module, qui peut être modifié ou même échangé indépendamment des autres. Le compilateur ESTEREL a généré à partir de cette spécification modulaire une implémentation intégrée et efficace.

Enfin une troisième expérience a consisté à générer automatiquement un système de communication à partir des spécifications des contraintes de communication d'une application de type client/serveur. ESTEREL a été utilisé pour décrire la partie contrôle de l'application et ASN.1 pour décrire les structures de données échangées à travers le réseau.

3.3.2 Expériences avec ALF

Participants : Isabelle Chrisment, Christian Huitema, Ellen Siegel

L'idée de base d'Application Level Framing (ALF) est de faire en sorte que ce soit le protocole qui s'adapte aux besoins de l'application et non l'inverse. L'application est considérée comme la plus apte à déterminer les stratégies à adopter lorsque des données sont perdues ou arrivent hors séquence. Cette conception d'une architecture intégrée et flexible nécessite d'une part que le système de communication ait accès à la sémantique de l'application et d'autre part que l'application ait aussi

un moyen d'agir sur les aspects de contrôle et de synchronisation du système de communication.

Afin de valider cette approche intégrée, nous avons voulu analyser et mesurer quel était le véritable impact d'ALF sur l'architecture et les performances du système de communication.

Pour notre étude, nous avons développé une application de transfert d'images JPEG qui se prête bien au concept ALF. Notre application utilise le format JFIF (JPEG File Interchange Format). JFIF est un format minimal de fichier qui permet d'échanger des images compressées JPEG.

Notre analyse porte principalement sur les points suivants:

- la taille de l'ADU

L'ADU est l'unité de données au niveau applicatif. Pour que ALF ait vraiment un sens, la taille de l'ADU doit être préservée à travers tout le système de communication et la fragmentation évitée. Ainsi une étroite relation doit être mise en place entre l'ADU et la MTU minimale (MTU: taille maximale d'un paquet sans fragmentation à travers le réseau).

- le traitement hors séquence

Traiter les paquets hors séquence nous paraît un facteur important d'accroissement des performances. Les unités de données peuvent en effet être délivrées directement à l'application sans attente du paquet suivant ou précédent.

Des mesures de performance sont actuellement en cours de réalisation. Mais déjà les résultats préliminaires montre l'intérêt de l'approche ALF.

3.3.3 Expériences avec ILP

Participants: Torsten Braun, Claude Castelluccia, Walid Dabbous, Christophe Diot

Nous avons effectué plusieurs expérimentations afin de valider le principe ILP. En particulier, une application de transfert de fichier selon le modèle client serveur a été implémentée. Le module de communication développé manuellement intègre le codage de présentation, une fonction de chiffrement simplifiée, ainsi qu'une implémentation de TCP au niveau utilisateur. Cette intégration a permis d'optimiser le nombre des accès

mémoire. Les résultats préliminaires des tests de performance ont montré que l'implémentation intégrée permet d'obtenir un gain de performance important par rapport à l'implémentation en couche. Nous sommes en train d'effectuer un étude plus détaillée des gains de performance escomptés avec ILP. En 1995, nous allons participer à la conception du compilateur permettant de générer automatiquement des implémentations ILP. Nous allons, entre autre, étudier l'impact de ILP sur les spécifications de protocoles en plus de détail.

3.3.4 Compilateur de Protocole

Participants : Isabelle Chrisment, Christophe Diot, Anthony Richards

Nous travaillons actuellement à la conception d'un premier prototype de la partie frontale du compilateur. Nous utilisons pour cela l'application de transfert d'images JPEG. A partir de la description de cette application en ESTEREL, le support de communication le plus adapté est sélectionné puis intégré à la spécification de l'application avant de produire automatiquement une implantation performante. Dans un premier temps nous prenons en compte dans le prototype uniquement les aspects ALF. Nous envisageons ensuite d'étendre notre prototype à d'autres applications et d'intégrer les travaux sur ILP au niveau du compilateur pour la fin 1995.

4 Actions industrielles

4.1 PASSWORD

Le projet VALUE PASSWORD a pour objet l'interconnexion et l'expérimentation de différents systèmes de sécurité, en particulier les systèmes basés sur les standards RSA et X509. Le projet regroupe trois consortiums: UCL + XTEL en Grande-Bretagne, GMD + DANET en Allemagne, et INRIA + E3X en France.

4.2 Collaboration avec Dassault Electronique

La collaboration avec Dassault Electronique et la DRET s'est renforcée cette année avec la signature d'un contrat de recherche et l'obtention d'un financement de thèse sur trois ans. Le sujet de cette thèse est "Modèle de Communication pour les Applications Temps-réel". Le propos de

cette thèse est dans un premier temps d'identifier les besoins en termes de services des applications temps-réel à fortes contraintes temporelles, et de définir les mécanismes de communication qui permettent de fournir les services identifiés. Une fois cette étape terminée, l'étudiant devra réaliser une application temps-réel dont les besoins sont ceux qui ont été définis dans la phase précédente de sa thèse. Dans une troisième étape, il devra concevoir un support de communication qui permette à l'application temps-réel qu'il a conçue de transmettre ses informations dans les contraintes qui sont celles d'une telle application. Ce support de communication devra intégrer les concepts nouveaux d'implantation hautes performances tels que ALF et ILP. La réalisation de ce système de communication temps réel se terminera par une évaluation quantitative et qualitative en environnement multimédia. Afin de permettre l'intégration des résultats de ces travaux dans les outils de production automatique d'implantation hautes performances qui sont en cours de développement dans HIPPARCH, les besoins de l'application temps-réel, ainsi que les mécanismes de communication associés seront spécifiés en ESTEREL.

4.3 Collaborations avec Thomson CSF

Nous étudions les problèmes associés à l'interconnexion de simulateurs militaires distribués sur un réseau étendu tel le réseau interarmées Socrate. Le but de l'étude est de proposer des architectures de réseau à court, moyen, et long terme qui offrent les services nécessaires (en particulier la transmission multipoint et la garantie de délais de transmission très faibles) pour permettre des simulations à grande échelle réalistes.

Une autre collaboration se met en place actuellement pour étudier des mécanismes de transmission multipoint fiable.

4.4 Collaborations avec le CNET

Une première étude effectuée en collaboration avec le projet MISTRAL a pour but de développer des modèles analytiques permettant d'évaluer la performance du protocole TCP sur un réseau étendu de type RENATER. Cette étude d'un an se terminera à la mi 1995.

Une deuxième étude sur 3 ans commence bientôt. Elle a pour but de comprendre les caractéristiques du trafic dans les réseaux, et d'utiliser les

résultats obtenus pour développer de meilleurs mécanismes d'allocation de ressources.

5 Actions nationales et internationales

5.1 Actions nationales

5.1.1 GDR/PRS et Action Réseau

Au plan national, RODEO participe à deux projets nationaux (en collaboration avec le LAAS, le MASI, et l'IRISA) dans le cadre du GDR/PRC et de "l'Action Réseau" du CNRS. Ces deux projets concernent respectivement l'architecture de protocole haute performance et la transmission multipoint fiable.

5.1.2 BULL

Nous travaillons aussi en collaboration avec BULL Grenoble sur des architectures de communication hautes performances en environnement Streams/UNIX. Un thésard commence actuellement sa troisième année de thèse chez BULL sous la direction de Christophe Diot.

5.2 Actions internationales

5.2.1 MICE

Nous participons à la deuxième phase du projet Européen MICE (Multimedia Integrated Conferencing for Europe) qui a pour but la définition et le déploiement de standards de vidéoconférence pour les réseaux de recherche. Nos partenaires sont UCL, SICS, GMD, les universités d'Oslo (UiO), de Bruxelles (ULB) et de Stuttgart (RUS). Après la première phase qui visait à avoir une plate-forme commune et qui s'est terminée en décembre 1993, nous nous intéressons actuellement dans le cadre de ce projet au support des applications de travail collaboratif sur les réseaux à haut débit.

5.2.2 HIPPARCH

Nous participons aussi à une collaboration internationale entre l'Europe et l'Australie sur le thème d'architecture flexible des protocoles haute

performance. HIPPARCH a pour but l'étude des techniques de spécification et de compilation de protocoles de communication afin de produire automatiquement des implantations performantes adaptées aux besoins de l'application. Ce projet est financé par la CEE; il s'effectue en partenariat avec UCL (Grande Bretagne), SICS (Suède), et UTS (Australie). La durée de ce projet est de deux ans à partir de Janvier 1994.

5.2.3 Activités dans le cadre de l'Internet

Christian Huitema a été élu président de l'*Internet Activity Board* (IAB) en Avril 1993. L'IAB est responsable de l'évolution de l'Internet, du point de vue technique. Il supervise le travail de l'*Internet Engineering Task Force* (IETF), et se préoccupe de l'intégrité architecturale du système et de ses évolutions pour faire face aux accroissements exponentiels de taille, vitesse et complexité du réseau.

Nous participons aux études menées par l'IETF dans un certain nombre de domaines. En particulier, pour définir la nouvelle version du protocole IP, de nouveaux protocoles pour les applications multimédia (le protocole RTP) et un mécanisme de contrôle de session de vidéoconférence.

5.2.4 Organisation de conférences

Le projet RODEO a organisé en Décembre 1994 le premier workshop HIPPARCH '94, dont le but est de réunir les équipes de recherche qui s'intéressent aux problèmes des architectures nouvelles, et en particulier aux concepts tels qu'ALF et ILP.

Le projet RODEO a aussi co-organisé cette année la conférence IFIP internationale "High Performance Networking" (HPN '94) qui a eu lieu à Grenoble fin Juin 1994. Cette conférence a rencontré un très vif succès avec près de 180 participants.

5.2.5 COST 237

Christophe Diot est représentant Français dans le projet Européen COST 237: Multimedia Telecommunication Services. Ce nouveau projet compte établir un forum Européen où seront discutés les évolutions des supports de communication Multimedia.

5.2.6 Comités de programme

Christian Huitema est membre des comités de programme des Congrès UPLAA '94 (Barcelone, Juin 1994) et SIGCOMM '94 (Londres, Septembre 1994).

Walid Dabbous est membre des comités de programme du congrès HPDC-3 (San Francisco, Août 1994) et HPDC-4 (Washington, Août 1995).

Jean-Chrysostome Bolot est membre des comités de programme des congrès Infocom '94 (Toronto, Juin 1994) et SIGCOMM '95 (Massachusetts, Août 1995).

Christophe Diot est membre du comité de programme du congrès HPN '94 (Grenoble, Juin 1994).

6 Diffusion des résultats

Une partie importante des résultats obtenus est diffusée dans le cadre des actions industrielles et internationales et par la mise à disposition de la collectivité des logiciels développés (M.Plus, Mavros, Thorn).

6.1 Démonstration de Vidéoconférence

En 1994, le projet RODEO a organisé plusieurs démonstrations des outils de travail collaboratif sur réseau INTERNET dont en juillet 1994:

- le 8 pour Jean-Marie Cavada, la marche du siècle, France 3.
- le 11 pour une délégation du Sénat, de la SFP, du CSA et de l'Education Nationale.
- le 20 pour Mme Ferrier du Cabinet du Premier Ministre accompagnée par Mr le Sénateur Laffite et Mr Ferrari le Recteur, Chancelier de l'Université de Nice Sophia Antipolis.
- le 25 pour MM. Pelchat, Député de l'Essonne Secrétaire de l'OPEST (Office Parlementaire pour l'Evaluation des choix Scientifiques et Techniques), Fleurieu, Conseiller au Sénat (OPEST), Grimond, Directeur Général Ellipse (la société de production pour Canal +), Crosse, Délégué général pour la Route des Haute Technologies.

Le logiciel IVS a été aussi utilisé pour la transmission de manifestations scientifiques sur le réseau Internet dont: Des réunions de travail par

vidéoconférence hebdomadaires ainsi que des séminaires mensuels dans le cadre du projet MICE. De même, nos résultats ont été présentés au grand public lors des journées *Sciences en fête* organisées par la demande du ministère de la recherche.

L'outil de tableau blanc Mscrawl a été aussi testé dans ce cadre.

6.2 Diffusion de produits

6.2.1 M.PLUS, PIZARRO

Les produits M.PLUS (système de messagerie OSI) et PIZARRO (annuaire OSI) sont commercialisés par E3X/SYNC.

6.2.2 MAVROS

Le compilateur ASN.1 MAVROS fait l'objet d'une coopération avec SIEMENS, et d'accords de licence avec E3X et Bull.

6.3 Actions d'enseignement

6.3.1 Enseignement universitaire

Réseaux couches hautes cours ESSI par C. Huitema.

Algorithmique des couches basses cours ESSI et Mastère CERAM, par W. Dabbous.

Protocoles Haut Débit cours à l'Ecole d'été Réseaux, Télécom Paris par W. Dabbous.

Réseaux à haut débit cours de DEA, Université de Nice-Sophia Antipolis, par W. Dabbous, J. Bolot, C. Huitema.

Applications multimedia sur Internet cours EURECOM et Télécom Bretagne par J. Bolot.

6.4 Participation à des conférences et colloques

Les membres du projet ont participé à plusieurs conférences et workshops. Veuillez vous reporter à la bibliographie pour une liste complète.

7 Publications

Livres et monographies

- [1] J. CROWCROFT, C. HUITEMA (éd.), *High Performance Protocols Architecture*, Sophia Antipolis, France, 1994.
- [2] C. HUITEMA, *Le Routage dans l'INTERNET*, Eyrolles, 1994.

Articles et chapitres de livre

- [3] J. BOLOT, «Characterizing end-to-end packet delay and loss in the Internet», *Journal of High Speed Networks* 3, 2, 1994.
- [4] W. DABBOUS, «Analysis of a Delay Based Congestion Avoidance Algorithm», in: *The OSI95 Transport Service with Multimedia Support*, Springer Verlag, Berlin, 1994, p. 280–296.
- [5] C. DIOT, W. DABBOUS, «Multimédia : Repenser le fonctionnement des protocoles», *Revue Normatique*, 54, Février 1994.
- [6] T. TURLETTI, «The INRIA Videoconferencing System (IVS)», *ConneXions - The Interoperability Report*, Octobre 1994, p. 20–24.

Communications à des congrès, colloques, etc.

- [7] H. AFIFI, J. BOLOT, «Using cache partitionning in X.500 for faster name resolution in X.400», in: *Proceedings of the Intl. Conference on Upper Layers (ULPAA '94)*, Barcelone, Espagne, Juin 1994.
- [8] U. BILTING, A. SASSE, C.-D. SCHULZ, T. TURLETTI, «International Research Seminars through Multimedia Conferencing: Experiences from the MICE project», in: *Proceedings of the Broadband Islands Conference*, Hamburg, Juin 1994.
- [9] J. BOLOT, T. TURLETTI, I. WAKEMAN, «Scalable feedback control for multicast video distribution in the Internet», in: *Proceedings of ACM Sigcomm '94*, p. 58–67, Londres, UK, Septembre 1994.
- [10] J. BOLOT, T. TURLETTI, «A rate control scheme for packet video in the Internet», in: *Proceedings of IEEE Infocom '94*, p. 1216–1223, Toronto, Canada, Juin 1994.
- [11] C. CASTELLUCCIA, W. DABBOUS, «Modular Communication Subsystem Implementation using a Synchronous Approach», in: *Proceedings of the Usenix Symposium on High Speed Networks*, Oakland, USA, Août 1994.
- [12] I. CHRISMENT, C. HUITEMA, «A Remote Operations System Tailored to Application Requirements», in: *Proceedings of the IFIP Conference*

- on Upper Layer Protocols, Architectures and Applications 94*, Barcelone, Espagne, Juin 1994.
- [13] I. CHRISMENT, «Impact of ALF on Communication Subsystems Design and Performance», *in: first HIPPARCH Workshop*, INRIA Sophia Antipolis, Sophia Antipolis, Décembre 1994. available on the INRIA anonymous ftp server.
 - [14] W. DABBOUS, «High performance presentation and transport mechanisms for integrated communication subsystems», *in: Proceedings of the 4th International IFIP Workshop on Protocols for High Speed Networks*, Vancouver, Canada, Août 1994.
 - [15] W. DABBOUS, «Protocoles Haut débit», *in: Actes de l'École d'Été Réseaux de Communication et Techniques Formelles*, Télécom Paris, Septembre 1994.
 - [16] C. DIOT, R. DE SIMONE, C. HUITEMA, «Communication Protocol Development Using ESTEREL», *in: first HIPPARCH Workshop*, INRIA Sophia Antipolis, Sophia Antipolis, Décembre 1994. disponible en ftp anonyme sur zenon.inria.fr.
 - [17] C. DIOT, R. DE SIMONE, C. HUITEMA, «Communication Protocol Engineering Using ESTEREL», *in: Colloque Informatique Communicante et Systèmes Distribués*, IMAG, Grenoble, Décembre 1994. papier invité.
 - [18] C. DIOT, «Reliability in Multicast Services and Protocols ; a Survey», *in: International Conference on Local and Metropolitan Communication Systems*, IFIP, Chapman and Hall, Kyoto, Décembre 1994.
 - [19] P. HOSCHKA, C. HUITEMA, «Automatic Generation Of Optimised Code for Marshaling Routines», *in: IFIP International Working Conference on Upper Layer Protocols, Architectures and Applications*, IFIP, North-Holland, p. 131–146, Barcelona, Spain, 1994.
 - [20] V. ROCA, C. DIOT, «A High Performance Streams-based Architecture for Communication Subsystems», *in: International IFIP Workshop on Protocol for High Speed Networks*, IFIP, Elsevier-North Holland, Vancouver, Août 1994.
 - [21] A. SASSE, U. BILTING, C.-D. SCHULZ, T. TURLETTI, «Remote Seminars through MultiMedia Conferencing: Experiences from the MICE project», *in: Proceedings of the INET'94/JENC5*, p. 251/1–251/8, Prague, Juin 1994.
 - [22] T. TURLETTI, J. BOLOT, «Issues with multicast video distribution in heterogeneous networks», *in: Proceedings of 6th Packet Video Workshop*, p. F3.1–F3.4, Portland, OR, Septembre 1994.
 - [23] T. TURLETTI, «A congestion control for packet video in the Internet», *in: INTEROP 94, Conference Note*, INTEROP, Paris, Octobre 1994.

Rapports de recherche et publications internes

- [24] C. CASTELLUCIA, I. CHRISMENT, W. DABBOUS, C. DIOT, C. HUITEMA, E. SIEGEL, R. DE SIMONE., « Tailored Protocol Development Using ESTEREL », *Rapport de Recherche n°2374*, Inria, Octobre 1994.
- [25] C. DIOT, « Comment Spécifier un Protocole ou un Ensemble de Protocoles en ESTEREL », *Rapport interne-projet rodeo*, INRIA, Octobre 1994.
- [26] T. TURLETTI, C. HUITEMA, « Packetization of H.261 video streams », *Internet draft*, Inria, Septembre 1994.

8 Abstract

The main objective of the project is to design, implement, and evaluate mechanisms that will allow the efficient and flexible use of high-speed networks (with a bandwidth greater than 1 Gb/s).

To achieve the above objective, we focus our activities in 2 main areas, namely

- Transmission control protocols for high-speed networks.
We are designing and developing flexible and high performance transmission control protocols which integrate mechanisms for the allocation of resources in the network, for multicast delivery, for flexible error control, etc. These protocols are used in particular in IVS (INRIA Videoconference System), an application for videoconference over the Internet.
- High performance applications.
Previous research in the RODEO project led to the development of the MAVROS compiler, and of fast algorithms for the management of data transparency in heterogeneous systems. We are extending this work so as to allow the compiler to handle detailed specifications of the applications. This will lead to further optimizations.

We expect that the two research areas described above will eventually merge within the framework of the recently-proposed Application Layer Framing (ALF) architecture. It will then be possible to take into account specifications of an application, of the network resources required by the application, and to automatically generate tailored communica-

tion modules that handle both data transparency, synchronization, and transmission control for this application.

Table des matières

1	Composition de l'équipe	1
2	Présentation du projet	2
3	Actions de recherche	4
3.1	Contrôle de transmission	4
3.1.1	Analyse de mécanismes de contrôle	4
3.1.2	Contrôle de transmission pour applications multi-media sur Internet	5
3.1.3	Tableau blanc et transmission multipoint fiable	9
3.1.4	Contrôle et garanties de qualité de service	11
3.2	Codage de Présentation à haut débit	13
3.3	Architecture flexible de protocole haute performance	16
3.3.1	Spécification de protocoles avec ESTEREL	17
3.3.2	Expériences avec ALF	18
3.3.3	Expériences avec ILP	19
3.3.4	Compilateur de Protocole	20
4	Actions industrielles	20
4.1	PASSWORD	20
4.2	Collaboration avec Dassault Electronique	20
4.3	Collaborations avec Thomson CSF	21
4.4	Collaborations avec le CNET	21
5	Actions nationales et internationales	22
5.1	Actions nationales	22
5.1.1	GDR/PRS et Action Réseau	22
5.1.2	BULL	22
5.2	Actions internationales	22
5.2.1	MICE	22

5.2.2	HIPPARCH	22
5.2.3	Activités dans le cadre de l'Internet	23
5.2.4	Organisation de conférences	23
5.2.5	COST 237	23
5.2.6	Comités de programme	24
6	Diffusion des résultats	24
6.1	Démonstration de Vidéoconférence	24
6.2	Diffusion de produits	25
6.2.1	M.PLUS, PIZARRO	25
6.2.2	MAVROS	25
6.3	Actions d'enseignement	25
6.3.1	Enseignement universitaire	25
6.4	Participation à des conférences et colloques	25
7	Publications	26
8	Abstract	28