

Projet ReMaP

Régularité et parallélisme massif

ENS-Lyon, CNRS et INRIA Rhône-Alpes

THÈME 1A



*R*apport
*d'**A*ctivité

2000

Table des matières

| | | |
|----------|---|-----------|
| 1 | Composition de l'équipe | 3 |
| 2 | Présentation et objectifs généraux | 4 |
| 3 | Fondements scientifiques | 5 |
| 3.1 | Algorithmique, bibliothèques et compilateurs | 7 |
| 3.2 | Support exécutif pour le métacomputing | 9 |
| 4 | Domaines d'applications | 11 |
| 5 | Logiciels | 13 |
| 5.1 | Outils pour le calcul parallèle scientifique | 13 |
| 5.2 | Environnements d'exécution haute performance | 14 |
| 5.3 | Applications pour le multimédia | 16 |
| 6 | Résultats nouveaux | 17 |
| 6.1 | Algorithmique, bibliothèques et compilateurs | 17 |
| 6.1.1 | Algorithmique sur réseau hétérogène de processeurs | 17 |
| 6.1.2 | Algorithmique des télécommunications | 19 |
| 6.1.3 | Transformation de programmes | 22 |
| 6.1.4 | Serveurs de calculs et optimisation d'algorithmes numériques parallèles | 23 |
| 6.2 | Environnements d'exécution parallèles et distribués | 25 |
| 6.2.1 | Environnement PM2 | 25 |
| 6.2.2 | Communications multi-protocoles | 26 |
| 6.2.3 | Évolution de PM2 pour le support du langage Java | 27 |
| 6.2.4 | MPI au-dessus de PM2 | 28 |
| 6.2.5 | Évolutions de BIP | 28 |
| 6.2.6 | Activités multimédia | 29 |
| 7 | Contrats industriels (nationaux, européens et internationaux) | 29 |
| 7.1 | <i>LHPC</i> | 29 |
| 7.2 | Collaboration avec Myricom | 31 |
| 7.3 | Contrat INRIA/Microsoft | 31 |
| 7.4 | Accord-cadre INRIA avec la société Alcatel | 31 |
| 8 | Actions régionales, nationales et internationales | 32 |
| 8.1 | Actions nationales | 32 |
| 8.2 | Actions financées par la Commission européenne | 33 |
| 8.3 | Réseaux et groupes de travail internationaux | 33 |
| 8.4 | Relations bilatérales internationales | 34 |

| | | |
|-----------|--|-----------|
| 9 | Diffusion de résultats | 35 |
| 9.1 | Animation de la communauté scientifique | 35 |
| 9.2 | Enseignement universitaire | 36 |
| 9.3 | Autres enseignements | 37 |
| 9.4 | Participation à des colloques, séminaires, invitations | 38 |
| 10 | Bibliographie | 39 |

Le projet ReMaP est un projet commun CNRS/ENS Lyon/INRIA du Laboratoire de l'informatique du parallélisme (LIP), UMR CNRS/ENS Lyon/INRIA 5668. Ce projet est localisé à Lyon dans les locaux de l'ENS Lyon.

1 Composition de l'équipe

Responsable scientifique

Yves Robert [Professeur ENS Lyon et UCB Lyon (jusqu'au 31/08/00)]

Frédéric Desprez [CR INRIA (à partir du 01/09/00)]

Assistants de projet

Sylvie Boyer [Contractuelle, 30% sur le projet]

Anne-Pascale Botonnet [Adjointe administrative ENS Lyon, 30% sur le projet]

Personnel Inria

Frédéric Desprez [CR]

Jean-François Méhaut [CR, détaché de l'univ. Lille 1 depuis le 01/10/98, départ le 01/09/00 (Prof. univ. Antille-Guyanne)]

Isabelle Guérin-Lassous [CR, arrivée le 01/09/00]

Personnel CNRS

Alain Darte [CR]

Nicolas Schabanel [CR, arrivé le 01/09/00]

Jean-Christophe Mignot [IR]

Loïc Prylli [CR]

Personnel ENS Lyon

Olivier Beaumont [Maître de conférences]

Luc Bougé [Professeur]

Stéphane Domas [ATER à 50%, départ le 01/09/00 (Mdc IUT Belfort)]

Raymond Namyst [Maître de conférences]

Ingénieurs experts

Laurent Bobelin [IE Inria, départ le 31/10/00]

Patrick Dargent [Ingénieur contractuel ENS Lyon, 3 mois]

Chercheurs extérieurs

Jean-François Méhaut [Prof. Univ. Antille-Guyanne (à partir du 01/09/2000)]

Chercheurs doctorants

Gabriel Antoniu [Allocataire MENRT]

Olivier Aumage [Allocataire MENRT]

Alice Bonhomme [Boursière Cifre *MS&I*]

Vincent Boudet [Allocataire moniteur normalien]

Frédérique Chaussumier [Boursière Cifre *MS&I*]

Vincent Danjean [Allocataire moniteur normalien, arrivée le 01/09/00]

Guillaume Huard [Allocataire MENRT]

Arnaud Legrand [Allocataire moniteur normalien, arrivée le 01/09/00]

Guillaume Mercier [Allocataire MENRT, arrivée le 01/09/00]

Christian Perez [Allocataire moniteur normalien, départ le 01/09/00 (CR Inria Rennes)]

Martin Quinson [Allocataire MENRT, arrivée le 01/09/00]

Fabrice Rastello [Allocataire MENRT, départ le 01/09/00 (Ingénieur ST Microelectronics)]

Frédéric Suter [Allocataire MENRT]

2 Présentation et objectifs généraux

Mots clés : data-parallélisme, parallélisation automatique, compilation, environnement de programmation, bibliothèques, réseaux à haut débit, application répartie, algorithmique hétérogène, métacomputing, télécommunications.

À sa création, l'objectif du projet *ReMaP* était de contribuer à l'élaboration des connaissances dans le domaine du calcul massivement parallèle régulier. La compilation data-parallèle reste au cœur du projet, mais nous nous intéressons aussi aux applications sur

réseaux hétérogènes (bibliothèques, processus légers, protocoles de communication et leurs impacts) et aux télécommunications sur réseaux sans fil (protocoles).

Les axes de recherche de *ReMaP* sont les suivants :

- techniques de parallélisation automatique et outils de parallélisation ;
- outils pour l’algèbre linéaire parallèle : algorithmique parallèle en algèbre linéaire dense et creuse, mise en œuvre de bibliothèques numériques sur réseaux hétérogènes de processeurs ;
- protocoles pour les réseaux sans fil ;
- compilation data-parallèle pour processus légers migrables : HPF (High Performance Fortran), C* et Java, support d’exécution multi-threads ;
- protocoles orientés applications pour les réseaux à haut débit.

Un point fort du projet est son ancrage industriel et ses activités de transfert. Citons principalement :

- laboratoire *LHPC* commun avec MATRA Systèmes et Information (*MS&I*) ;
- collaboration avec *MS&I*, le centre multimédia *Érasme*, *Rhône Vision Câble* et *Tonna Informatique* dans le cadre du projet *CHARM* ;
- collaboration avec *Myricom* autour des développements de *BIP* ;
- coopération avec l’université du Tennessee, Knoxville, dans le cadre de la bibliothèque d’algèbre linéaire parallèle *ScaLAPACK* qui est mise à la disposition de la communauté internationale.

Une vision schématique mais synthétique de nos axes de recherches est proposée à la Table 1. Les mots-clés de *ReMaP* sont :

| |
|---|
| <i>Compilation + Bibliothèques + Réseaux + Applications</i> |
|---|

3 Fondements scientifiques

Depuis quelques années, le paysage des architectures parallèles s’est considérablement modifié, et ce aux deux extrémités du spectre : d’une part le parallélisme est descendu au niveau interne des processeurs (superscalaire, VLIW), et d’autre part on assiste à l’avènement des grappes de PC (*cluster computing*), avec l’interconnexion à plus large échelle de ressources de calcul distribuées (*métacomputing*). Un des nouveaux défis qui se pose pour la mise en œuvre des applications parallèles est la maîtrise de l’hétérogénéité. Pour les utilisateurs et les programmeurs, le caractère hétérogène des architectures pose de nouveaux

| Niveau | Logiciel | Collaborations | |
|-------------------------|---------------------------------------|--|--|
| | | France | Internat. |
| Architectures | Piles de PC, serveur multiméda | MS&I | |
| Protocoles | BIP, SCI | RHDM, VASY, ARC ResCapA | Myricom, UC Berkeley |
| Bibliothèques | MPI, PVM, GCM ScaLAPACK | LORIA | U Tenn., Argonne |
| Supports d'exécution | PM2, Dosmos | LIFL, Apache, Sirac, Caps | Argonne, Rice |
| Compilateurs | HPF, C* | ex Paradigme/PRS, iHPerf | GMD, U New Hamp. Rice, HP Labs (CA) |
| Environnements | TransTool, Scilab, Nestor, PlusPyr | Métalau, Résédas, LaBRI LIFC, LaRIA | |
| Applications | Numérique dense + creux Cache Web | LaBRI MS&I | TTN |

TAB. 1 – *Vision synthétique de ReMaP*

problèmes dans le développement des applications. Deux approches pour la conception des applications parallèles sont possibles.

- L'application est conçue de manière séquentielle et le programmeur doit pouvoir disposer d'outils de parallélisation automatique et d'interfaces conviviales. Cette problématique est rendue particulièrement complexe avec les nouvelles technologies de processeurs qui permettent par exemple l'exécution simultanée de plusieurs instructions. Les compilateurs de langages, même séquentiels, doivent maintenant réussir à extraire du parallélisme des programmes.
- L'application est conçue de manière parallèle et de nombreuses difficultés surgissent : difficultés algorithmiques bien sûr, mais aussi difficultés de mise au point qui dépendent de la complexité d'une plate-forme parallèle hétérogène. Les principaux problèmes concernent les communications (différentes technologies de réseaux et de protocoles), la répartition des données qui tient compte de l'hétérogénéité (puissance de calcul, mémoire, bande passante) et répartition des calculs.

L'objectif de *ReMaP* est de (tenter de) relever ces nouveaux défis, en apportant des contributions au niveau des algorithmes, des bibliothèques, de la compilation, des environnements de programmation et des protocoles pour les réseaux à haut débit. La pyramide des problèmes auxquels nous nous intéressons a été décrite succinctement au paragraphe précédent ; ceux-ci seront abordés de manière plus précise ci-dessous.

Notre credo est triple :

- participer aux projets de collaboration internationaux (comme *ScaLAPACK*) plutôt que de développer en interne des logiciels propriétaires ;

- travailler en forte collaboration avec des partenaires industriels pour valider notre approche et nos résultats ;
- mener de front des travaux de recherche fondamentale et appliquée.

Ces objectifs, bien que parfois difficilement conciliables, guident nos travaux. Ceux-ci sont structurés en deux grands axes :

1. Algorithmique, bibliothèques et compilateurs ;
2. Environnement d'exécution multi-threads et réseaux à haut débit.

3.1 Algorithmique, bibliothèques et compilateurs

Participants : Olivier Beaumont, Vincent Boudet, Frédérique Chaussumier, Alain Darte, Frédéric Desprez, Stéphane Domas, Isabelle Guérin-Lassous, Guillaume Huard, Arnaud Legrand, Martin Quinson, Fabrice Rastello, Frédéric Suter, Nicolas Schabanel.

L'objectif de ce premier thème du projet est de rendre le parallélisme *transparent* pour l'utilisateur, ou du moins de faciliter sa mise en œuvre.

Parallélisation automatique et transformation de programmes Il s'agit de développer et d'intégrer de nouvelles stratégies permettant de transformer (semi-)automatiquement des portions de code séquentiel (principalement des boucles Fortran) en codes annotés par des directives de type HPF (High Performance Fortran) ou OpenMP. Le but est d'aider le programmeur à identifier le parallélisme potentiel de son code au niveau des boucles et d'effectuer automatiquement les transformations nécessaires à sa place (ordonnancement, placement, partitionnement, etc.)

Si les technologies de processeurs évoluent rapidement, elles exigent des compilateurs de plus en plus performants pour les exploiter. Les processeurs actuels disposent aujourd'hui de capacités à exécuter simultanément plusieurs instructions par un pipeline logiciel. Les techniques de parallélisation automatique étudiées depuis plusieurs années devraient être (ou sont?) utilisées par les prochaines générations de compilateurs. Un des prochains enjeux concernera certainement la synthèse de circuits pour intégrer au niveau du matériel des dispositifs facilitant la conception des compilateurs.

Nos travaux présentent deux facettes, fortement couplées :

- développements théoriques relatifs aux problèmes de transformations de code, de leur classification théorique à la mise au point d'algorithmes les résolvant et à l'étude de la faisabilité de leur intégration au sein d'un compilateur-paralléliseur ;
- développements logiciels (*TransTool* et *Nestor*) dont le but est de mettre en œuvre et de valider nos techniques dans le cadre d'une plate-forme de parallélisation (semi-) automatique de programmes Fortran vers HPF. La complexité de la génération du code est un paramètre essentiel pour le choix des transformations source-à-source que nous étudions.

Bibliothèques et algorithmique parallèle hétérogène

Bibliothèque ScaLAPACK Nous contribuons à plusieurs titres au développement de la bibliothèque d'algèbre linéaire *ScaLAPACK* pour machines parallèles qui vise aussi bien la classe des supercalculateurs à mémoire partagée que celle des réseaux de stations de travail à mémoire distribuée :

- optimisation d'enchaînement pipeline de procédures de base d'algèbre linéaire ;
- recherche de nouveaux algorithmes de redistribution de tableaux, et amélioration des procédures existantes ;
- propositions algorithmiques d'extension des noyaux de calcul de *ScaLAPACK* aux grappes hétérogènes et aux collections de telles grappes.

Scilab parallèle *Scilab* est un logiciel de calcul scientifique de type Matlab développé par le projet Métalau à l'INRIA. Une intégration de PVM dans *Scilab* permet déjà l'accès à la bibliothèque *ScaLAPACK*. Nous développons une version parallèle de *Scilab* en collaboration avec les projets Métalau et Résédas de l'INRIA et les laboratoires LaBRI, LIFC et LaRIA. Ainsi l'utilisateur pourra-t-il appeler une procédure de la bibliothèque qui s'exécutera de manière relativement transparente sur les ressources parallèles auxquelles il a accès. Les problématiques de recherche derrière le développement de tels outils sont l'évaluation des performances statiques et dynamiques, l'équilibrage des charges, la recherche de distributions optimales, le traitement d'architectures hétérogènes de machines et l'algorithmique parallèle sur plateformes hétérogènes en général.

Algorithmique des télécommunications Cette activité vient de démarrer dans le cadre du recrutement d'Isabelle Guérin Lassous et de Nicolas Schabanel.

Il s'agit de développer des algorithmes pour résoudre différentes problématiques dans différents types de réseaux de télécommunications. Les problématiques abordées couvrent de nombreux aspects rencontrés sur ce type de réseaux puisqu'elles vont des problèmes de planification (placement, allocation de fréquences) jusqu'aux problèmes de qualité de service (réservation de bande passante, délai moyen, etc) en passant par les problèmes de routage (point-à-point ou/et multipoints) et de dissémination de données. Les réseaux de télécommunications étudiés concernent les réseaux filaires hétérogènes, les réseaux sans fil (réseaux satellitaires, réseaux cellulaires, réseaux locaux et réseaux ad hoc).

Nous essayons d'apporter nos connaissances en algorithmique séquentielle et distribuée afin de proposer des protocoles efficaces pour les problèmes énoncés ci-dessus. Nous pensons qu'aborder les différents aspects des réseaux de télécommunications (planification, routage, qualité de service) en même temps nous permet d'avoir une très bonne vue d'ensemble sur l'algorithmique des télécommunications d'une part, et que ces aspects sont très fortement liés d'autre part (par exemple, suivant la planification choisie, les protocoles de routage ou de mise en œuvre d'une certaine qualité de service peuvent être très différents).

Si, dans un premier temps, les résultats obtenus sont théoriques, il nous semble néanmoins très important de les valider par une approche expérimentale à l'aide de simulateurs.

3.2 Support exécutif pour le métacomputing

Participants : Gabriel Antoniu, Olivier Aumage, Alice Bonhomme, Luc Bougé, Vincent Danjean, Alexandre Denis, Jean-François Méhaut, Guillaume Mercier, Jean-Christophe Mignot, Raymond Namyst, Loïc Prylli.

Notre objectif: le métacomputing départemental Depuis quelques années, la notion de processus léger (*thread*) est apparue comme un outil très utile pour le calcul haute performance comme l'ont montré de nombreux projets de recherche de par le monde, notamment *Nexus* (Ian Foster, ANL) et *Chant* (Matthew Haines, ICASE). En France, on peut citer le système Athapascan 0 (projet Apache, Grenoble), contemporain de PM2.

D'autre part, l'apparition de réseaux à haut débit dédiés à la construction de plateformes parallèles (les grappes de stations) a eu un impact déterminant sur le monde du parallélisme durant la dernière décennie. En effet même si les concepts de base restent identiques aux générations précédentes de machines parallèles, ces réseaux ont rendu indépendants les choix des trois aspects principaux d'une plate-forme parallèle : les noeuds de calculs, le réseau de communication et le logiciel de base (système d'exploitation et services de communications). Parmi les travaux dans ce domaine de recherche, on peut citer Score/PM (RWCP, Japon), LFC (Vrije University, Hollande), NOW (Berkeley), Trapeze (Duke University), GAMMA (Université de Gênes, Italie).

Notre objectif de recherche est de conjuguer ces deux développements pour permettre l'utilisation de grappes de PC standards pour le calcul haute performance. Il s'agit donc d'une forme de *métacomputing* spécialisée pour le niveau *départemental*, par exemple à l'échelle d'un bâtiment ou d'un petit groupe de bâtiments physiquement proches les uns des autres. Cette proximité géographique permet de réduire considérablement les problèmes d'administration et de sécurité liés au métacomputing, sans toutefois les annuler totalement comme ce serait le cas pour une grappe privée. D'autre part, cette proximité permet de considérer une interconnexion à très haut débit, typiquement de l'ordre du gigabit par seconde, entre les noeuds. Dans ce cadre, les temps de calcul locaux et les temps de communication sont du même ordre de grandeur, et les interfaces entre les différents modules ont un impact direct sur les performances globales (copies superflues, en particulier). Le modèle d'architecture de référence est donc un réseau hiérarchique de grappes de PC (une *grappe de grappes*), chaque grappe étant interconnectée par un réseaux rapide (1 Gb/s). L'hétérogénéité des grappes est une donnée fondamentale du problème : hétérogénéité des processeurs (Intel, Alpha, etc., avec des cadencements variables), hétérogénéité des réseaux (Myrinet, SCI, Giganet, etc. pour les grappes, Fast-Ethernet pour l'interconnexion globale).

Notre approche se fonde sur deux développements logiciels étroitement coordonnés : PM2 et BIP/Myrinet.

L'environnement d'exécution multithread PM2 L'environnement multithread distribué PM2 (*Parallel Multithreaded Machine*) a été initialement développé au sein du projet Espace du LIFL, avant d'être transféré dans le projet ReMaP. PM2 est en fait constitué d'une *suite* de modules. Le module de gestion de thread, appelé *Marcel*, fournit une implémentation très

efficace des primitives POSIX en espace utilisateur. Le module de gestion des communications, appelé *Madeleine*, fournit une interface portable particulièrement optimisée pour les appels de procédures à distance (*Remote Procedure Call*, RPC). Le module de mémoire virtuellement partagée, appelé *DSM* pour *Distributed Shared Memory*, permet aux threads situés sur des nœuds différents de partager dynamiquement des pages de la mémoire virtuelle. Grâce à l'intégration de l'ensemble de ces modules, PM2 offre une fonctionnalité de *migration préemptive transparente* des threads entre nœuds distants. Cette fonctionnalité ouvre la possibilité d'installer au-dessus de PM2 un ordonnanceur global qui équilibre la charge dynamiquement en migrant les threads à travers l'ensemble du système.

L'ensemble de la conception de PM2 a été élaborée avec un souci de portabilité. La suite PM2 est actuellement disponible au-dessus de la plupart des systèmes : GNU/Linux, Solaris, ainsi que des variantes propriétaires (IBM et HP notamment). Le portage sur WindowsNT est en cours. L'interface de communication est disponible sur la plupart des interfaces utilisées pour l'interconnexion de grappes de calculateurs : TCP/IP bien sûr, mais aussi MPI, BIP pour le réseau gigabit Myrinet, SISI pour les réseaux SCI, etc. PM2 a en particulier été porté sur l'interface de communication VIA (*Virtual Interface Architecture*) actuellement en plein développement. Des évaluations de performance de VIA sur le réseau Gigaset sont en cours.

L'interface réseau BIP/Myrinet La suite logicielle BIP (*Basic Interface for Parallelism*) propose une interface optimisée pour le réseau *Myrinet* proposé par la société américaine *Myricom*. Cette interface est particulièrement adaptée au support d'environnement exécutifs multithreads tels que PM2 et/ou au support d'implémentation pour des bibliothèques de communication telles que MPI. La conception favorise les transmissions zéro-copies en espace utilisateur, ainsi que l'utilisation intensive des capacités de calcul et de gestion mémoire offertes par la carte Myrinet. Ceci permet de soulager d'autant le processeur principal et donc d'augmenter considérablement la réactivité et les performances. Ces travaux ont déjà permis d'obtenir d'excellents résultats pour l'utilisation efficace des réseaux Myrinet et plusieurs dizaines d'équipes dans le monde utilisent la suite logicielle BIP comme base pour leurs propres développements (BIP, MPI-BIP) ou comme support réseau performant pour leurs applications (MPI-BIP, IP-BIP).

Les grands défis de recherche

Communication multi-protocoles pour étendre les fonctionnalités de PM2, et tout particulièrement de son module de communication *Madeleine*, aux grappes de grappes constituées de sous-réseaux différents. L'objectif est d'obtenir ainsi des versions *portables* et *multi-protocoles* de couches de communication plus complexes, comme par exemple MPI.

Entrées-sorties distribuées, en vue de l'utilisation de PM2 pour l'implémentation de serveurs multimédia sur des grappes de PC standards. Nos expériences (caches Web, etc.) ont montré l'importance déterminante de mécanismes d'entrée-sortie *asynchrones* étroitement intégrés avec l'ordonnanceur de threads.

Support exécutif haute performance. Il s'agit ici d'utiliser PM2 comme cible pour des compilateurs. Il devient alors possible de faire coopérer des analyses statiques avec des ajustements dynamiques. Une première expérience a été conduite avec un compilateur HPF, avec des résultats concluants, et nous travaillons actuellement sur un compilateur Java. D'autre part, PM2 peut aussi offrir un support exécutif très intéressant pour des environnements complexes comme des ORB Corba ou des environnements de gestion comme Globus.

Optimisation des couches réseaux. L'efficacité de PM2 est entièrement dépendante de l'efficacité des couches réseaux sur lesquelles il s'appuie. L'optimisation fine de la bibliothèque de communication MPI au-dessus de l'interface BIP/Myrinet reste l'une de nos priorités à cause de son impact : plusieurs dizaines d'équipes l'utilisent à ce jour dans le monde. Parmi les travaux récents, citons le support optimisé des nœuds SMP et la mise en place de dispositifs de communication asynchrone au-dessus de BIP pour faciliter les recouvrement calcul et communications.

4 Domaines d'applications

Contexte Le parallélisme évolue et les domaines d'application se multiplient. Avec l'apparition de réseaux de stations de travail ou de piles de PC au bon rapport performance/prix, de nombreux utilisateurs se tournent vers des solutions parallèles. Outre les domaines d'application (on dépasse le traditionnel calcul scientifique pour aborder le secteur médical, bancaire, la sidérurgie, le textile, la publicité, la géographie, etc.), la nature des partenaires évolue (aux centres de recherche et développement des grands groupes publics ou privés s'ajoutent désormais les petites et moyennes entreprises). Ces nouveaux utilisateurs ont un besoin crucial d'environnements de programmation performants.

Calcul scientifique Nos objectifs se déclinent selon trois axes principaux.

- Concevoir et développer un environnement de programmation standard et portable (incluant des bibliothèques de calcul et de macro-communications, des outils de placement et d'ordonnancement, des compilateurs-paralléliseurs, etc.) pour machines parallèles à mémoire distribuée et réseaux de stations de travail.
- Mettre à disposition les plates-formes d'expérimentation développées en collaboration avec MS&I pour les chercheurs et les industriels ; offrir des services d'ingénierie pour le portage d'applications existantes ou le développement de nouvelles applications, et pour les activités de recherche associées.
- Assurer la formation des étudiants, chercheurs et ingénieurs et ainsi faciliter l'introduction du calcul parallèle, en particulier au niveau des services de recherche et développement des entreprises.

Applications aux télécommunications Nos objectifs sont d'appliquer nos connaissances en algorithmique parallèle et distribuée à l'algorithmique des télécommunications afin de proposer des protocoles de réseaux efficaces. Cette approche est menée selon deux axes.

- Fournir des algorithmes parallèles efficaces fonctionnant sur réseaux de stations de travail ou de piles de PC répondant à des problèmes de télécommunication (comme le placement d'équipements ou l'allocation de fréquences par exemple). Notre but est d'apporter des algorithmes plus efficaces (en terme de temps ou d'optimisation) que ceux existant jusqu'ici.
- Proposer des algorithmes distribués qui mettent en œuvre des protocoles réseaux. Ces algorithmes sont directement implantés sur les réseaux de télécommunications en question. Ils doivent donc prendre en compte les aspects architecturaux et protocolaires des réseaux. Dans le cas des réseaux mobiles, les algorithmes proposés doivent s'adapter à la topologie dynamique des réseaux.

Pour ce domaine d'applications qui démarre au sein du projet ReMaP, un premier projet avec des partenaires industriels vient de débiter (projet ProxiTV).

Applications multimédia Nous développons des logiciels permettant de transformer des grappes d'ordinateurs standards en serveurs de cache Internet haute performance pour les têtes de réseau des boucles haut débit. Ce système, réalisé conjointement par *MS&I* et le LIP, utilise des technologies issues du parallélisme pour garantir son extensibilité. Le système de cache intègre des fonctionnalités d'indexation en ligne et de filtrage. Un autre aspect novateur du système réside dans le support des flux audio et vidéo par des mécanismes de cache ou de miroir pour tenir compte de la part grandissante des documents multimédia dans le trafic Internet. Le projet inclut des phases d'expérimentation sur le réseau haut débit *Autoroutes Rhodaniennes de l'Information* de Rhône Vision Câble. Dans le cadre de ces expérimentations, *MS&I* et le LIP mettront un serveur de cache extrêmement performant à disposition du centre serveur Érasme. Cette plate-forme permettra de vérifier l'adéquation des solutions proposées en termes de fonctionnalités et de performances et d'évaluer les bénéfices des infrastructures à haut débit.

Un autre type d'application est étudié dans le cadre du projet SPIHD (Services et Programmes Interactifs pour l'Internet Haut Débit). Plus précisément, il s'agit de définir un ensemble de services de télévision interactive pour l'Internet haut débit. Quatre composantes ont été identifiées pour mener à bien ce projet.

1. La production du contenu multimédia par les partenaires de l'audio-visuel (France3-Lorraine, la cinquième, CanalWeb).
2. Le développement de l'infrastructure logicielle (numérisation, serveur de tête de réseau, gestion et mise à jour des bases de données).
3. L'expérimentation en grandeur réelle sur une boucle locale à Nancy avec la société SEM-Câble de l'Est.

4. L'évaluation technique et économique de la solution pour un éventuel déploiement à plus large échelle.

La contribution du projet *ReMaP* se situe au niveau du point 2 où nous nous intéresserons aux possibilités d'indexation de la base de données des vidéos. Un ingénieur-expert INRIA sera très prochainement engagé pour participer à ce projet et réaliser les développements logiciels.

5 Logiciels

5.1 Outils pour le calcul parallèle scientifique

Mots clés : Parallélisation de code, calcul numérique, parallélisation automatique, serveurs de calculs, bibliothèques numériques.

Participants : Alain Darte, Frédéric Desprez, Isabelle Guérin-Lassous, Guillaume Huard, Martin Quinson, Frédéric Suter.

Logiciel *Scilab* parallèle *Scilab* est un logiciel de calcul scientifique de type Matlab développé par le projet Métalau à l'INRIA. Nous parallélisons ce logiciel dans le cadre de l'ARC INRIA OURAGAN.

Nos développements concernent la mise en place de serveurs de calculs efficaces accessibles depuis l'interface Scilab et l'interfaçage directement dans Scilab de bibliothèques numériques telles que ScaLAPACK ou PETSc. Autour des serveurs de calculs, nous avons développé les prototypes de deux outils utilisés pour la localisation de ressources logicielles et matérielles adaptées à la résolution de problèmes envoyés depuis Scilab. Le premier outil est une base de données de ressources logicielles (SLiM, *Scientific Library Metaserver*). Le second outil permet l'évaluation des performances des calculs sur les serveurs et des communications entre le client et les serveurs, et les serveurs entre eux (FAST, *Fast Agent System Timer*).

Pour plus d'informations voir le site Web: <http://www.ens-lyon.fr/~desprez/OURAGAN/>.

Outil *Nestor* Le noyau du système de transformations de programmes *Nestor* terminé, ce sont ses fonctionnalités qui ont été enrichies. *Nestor* est un outil de manipulation d'arbres syntaxiques qui gère les entités élémentaires d'un code Fortran comme des objets C++. Ce logiciel (développé par Georges-André Silber), robuste, complètement documenté en HTML, mais simple d'utilisation grâce à son aspect orienté objet, sert de plate-forme d'implantation et de tests pour les algorithmes que nous étudions au niveau théorique. Tous nos algorithmes de transformations de code (voir la section 6.1.3) y sont implantés. Il sert également de noyau à aLaSca. Les nouvelles fonctionnalités de *Nestor* permettent notamment d'éviter la "fuite de mémoire" lors de l'utilisation des objets C++, de disposer de méthodes de parcours de l'arbre syntaxique, avant et arrière, de pouvoir implanter facilement de nouveaux algorithmes par

des mécanismes de *plug-in* qui évitent d'avoir à recompiler *Nestor* lors d'un nouvel ajout, etc. Georges-André Silber est maintenant maître-assistant aux Mines de Paris (ENSMP). Le développement de ce logiciel se fera donc à présent en collaboration avec l'ENSMP. Pour plus d'informations voir le site <http://www.cri.ensmp.fr/~silber/nestor/index.html>.

Logiciel Scheduler Scheduler est un petit logiciel d'évaluation d'heuristiques de pipeline logiciel qui nous a permis d'effectuer des tests de performance (virtuels, c'est-à-dire simulés) de nos techniques. Le logiciel, développé par Guillaume Huard, permet la spécification de machines simples de type VLIW, la génération aléatoire de graphes de dépendance, leur visualisation à l'aide de l'outil *vcg*, l'application des algorithmes de décalage (celui de Leiseron et Saxe et le nôtre), l'application de l'heuristique de Gasperoni et Schwiegelshohn (autre approche de pipeline logiciel décomposée), le calcul des bornes de performance classiques, ainsi que diverses compactations de boucle à base d'ordonnancement par liste.

Bibliothèque SSCRAP SSCRAP est une bibliothèque C++ qui met en œuvre les routines de base nécessaires pour les algorithmes écrits dans le modèle CGM (Coarse Grained Multicomputer) et fonctionnant sur plusieurs types d'architectures. Elle est dédiée au calcul parallèle sur les graphes. Cette bibliothèque est le fruit d'une collaboration avec Jens Gustedt (projet Résédas, INRIA Lorraine) et Mohamed Essaïdi (projet Résédas, INRIA Lorraine). La version actuelle est basée sur MPI. Les algorithmes implantés et disponibles sont le list ranking et la somme parallèle préfixe. D'ici la fin de l'année, le tri, les composantes connexes et la contraction d'arbre seront aussi disponibles. URL : <http://www.loria.fr/~gustedt/cgm/>.

5.2 Environnements d'exécution haute performance

Mots clés : Multithreading distribué, migration, RPC, réseaux à haut débit, mémoire partagée distribuée, Passage de messages, communication, réseaux à haut débit, Myrinet.

Participants : Gabriel Antoniu, Olivier Aumage, Alice Bonhomme, Luc Bougé, Jean-François Méhaut, Raymond Namyst, Loïc Prylli.

Environnement d'exécution PM2 PM2 est un environnement multithreads portable permettant d'exploiter efficacement les architectures distribuées haute performance (supercalculateurs, grappes de stations SMP interconnectées par réseau haut débit). Il se distingue par l'efficacité de sa gestion des processus légers, par ses fonctionnalités d'équilibrage dynamique de charge (migration de processus légers) ainsi que par son interfaçage efficace avec les protocoles de communication de très bas niveau. L'objectif de l'environnement PM2 est de définir un support d'exécution haute performance pour les applications parallèles ou les compilateurs de langages parallèles (HPF, Java) sur architectures distribuées (machines parallèles, grappes et bientôt grilles). La caractéristique majeure de ces applications est qu'il est difficile, voire impossible, de répartir statiquement les traitements et les données sur les processeurs de manière équilibrée.

Le modèle de programmation PM2 s'articule autour d'une décomposition des calculs en procédures activables par un mécanisme de type RPC (appel de procédure à distance). Ce découpage, potentiellement extrêmement fin, est pris en charge efficacement par le support exécutif de PM2. Afin de corriger les situations de déséquilibre, PM2 fournit un opérateur de migration permettant de déplacer les activités dynamiquement d'un processeur vers un autre. Cet opérateur s'appuie sur un mécanisme d'allocation iso-adresse assurant qu'un processus léger (et les données qu'il manipule) reste toujours logé dans la même zone d'adresses virtuelles des processus. En complément, PM2 inclut un noyau générique permettant la gestion d'une partie de l'espace mémoire en mode *virtuellement partagé* (DSM). Ce mécanisme permet aux processus légers d'une application PM2 de partager directement des données en mémoire sur des grappes homogènes.

Le support d'exécution s'appuie sur deux bibliothèques (Marcel et Madeleine) qui ont été développées pour PM2. Marcel est une bibliothèque de processus légers qui sont créés en contexte utilisateur (temps de commutation $< 1 \mu s$) et dont l'exécution est prise en charge par des processus noyaux permettant ainsi d'exploiter le parallélisme d'architectures de type SMP. Madeleine est une interface de communication qui se veut à la fois portable et efficace sur différents protocoles réseaux (temps de migration d'un processus léger : $26 \mu s$ sur le réseau SCI). Récemment, une version adaptative multi-protocoles de Madeleine a été conçue permettant l'utilisation de plusieurs protocoles réseaux au sein d'une même application. En outre, lorsqu'un protocole sous-jacent offre plusieurs modes de transfert des données, Madeleine sélectionne automatiquement le plus approprié de manière à rechercher le maximum d'efficacité. Cette interface est actuellement opérationnelle sur les protocoles BIP (Myrinet), SISI (SCI), MPI, et TCP.

Les codes source complets de la suite PM2 sont disponibles à l'URL <http://www.pm2.org>. L'environnement PM2 est en outre déposé à l'Agence de Protection des Programmes. Plusieurs équipes de recherches, en France et à l'étranger, utilisent PM2 ou certains de ses sous-composants : Rennes (Thierry Priol), Bordeaux (Jean Roman), UNH-Durham/USA (Phil Hatcher), Berlin/Allemagne (Frank Mueller).

Services de communications BIP BIP est un système logiciel fournissant plusieurs types de services de communications pour le réseau Myrinet.

La couche de plus bas niveau (firmware et bibliothèque hôte) s'interface directement avec le matériel. L'utilisation principale du système BIP se fait par l'intermédiaire de MPI-BIP, une implémentation complète du standard MPI-1 basée sur MPICH, permettant à la majeure partie des applications distribuées de tourner sur notre système sans effort de portage.

Plus de 200 sites distincts ont acquis un mot de passe pour rapatrier le système BIP. Les retours montrent que plusieurs dizaines d'équipes l'utilisent soit pour leurs propres développements de *middleware* (BIP, MPI-BIP), soit comme support réseau performant pour leurs applications (MPI-BIP, IP-BIP).

Les logiciels BIP sont actuellement disponibles pour l'ensemble des cartes Myrinet et pour les architectures x86/Alpha/PowerPc sous Linux.

Les nouvelles fonctionnalités de ce logiciels sont les suivantes :

- Réécriture des mécanismes d'enregistrement mémoire, modèle sécurisé et permettant

le `fork()`,

- support des grappes hétérogènes au niveau BIP et MPI-BIP,
- support d'une réception bloquante sans attente active, mais au contraire liée à l'ordonnanceur du système, (permettant donc d'avoir plus de processus que de processeurs).

5.3 Applications pour le multimédia

Mots clés : Système de fichiers distribués, tolérance aux pannes, serveur vidéo, Cache web distribué, grappes, serveur internet.

Participants : Alice Bonhomme, Luc Bougé, Jean-François Méhaut, Jean-Christophe Mignot, Loïc Prylli.

Système de stockage vidéo Alice Bonhomme dans le cadre d'une bourse CIFRE a réalisé avec la société Sycomore-Aérospatiale-Matra un système de stockage vidéo basé sur les principes suivants.

- Chaque flux est réparti sur un ensemble de nœuds avec des blocs de parité pour assurer la redondance nécessaire à la tolérance aux pannes.
- En cas de panne, il n'y a pas de perturbation des clients. Les blocs de parité sont lus à l'avance, ce qui permet l'alimentation du client avant qu'une panne soit confirmée.
- L'architecture du système est conçue pour être facilement extensible et il n'y a pas de nœud maître.

Ce système de stockage distribué est intégré dans des produits de cette société destinés à la commercialisation.

Cache Web parallèle Dans le cadre du projet *CHARM*, REMAP a réalisé un cache web parallèle *extensible* fondé sur une grappe de PC. Ce cache Web est en cours d'évaluation au centre ÉRASME. Le but est d'être extensible en terme de quantité de données cachées et de nombre de clients supportés.

Le composant principal de ce cache est un *routeur HTTP*, permettant de répartir les requêtes HTTP entre plusieurs unités de cache séquentiels suivant une fonction de partitionnement des URL.

L'utilisation des techniques de *VirtualServer* de Linux permet une première répartition *round-robin* des connexions TCP vers l'ensemble des *routeurs HTTP* (aussi appelés *proxy-light*) qui redistribuent ensuite les requêtes après analyse du contenu de l'en-tête HTTP.

Pour des raisons de performances, les communications entre le routeur HTTP et les unités de caches réparties sur les nœuds sont transmises via des messages MPI plutôt que sur une connexion TCP.

6 Résultats nouveaux

6.1 Algorithmique, bibliothèques et compilateurs

Participants : Olivier Beaumont, Vincent Boudet, Frédérique Chaussumier, Alain Darte, Frédéric Desprez, Isabelle Guérin Lassous, Arnaud Legrand, Stéphane Domas, Guillaume Huard, Martin Quinson, Fabrice Rastello, Yves Robert, Nicolas Schabanel, Frédéric Suter.

Mots clés : Algorithmique hétérogène, algorithmique des télécommunications, environnement de programmation, bibliothèque, parallélisation automatique.

Résumé : *Cette partie résume nos travaux récents concernant l'algorithmique, les bibliothèques et les compilateurs. Ces travaux sont à la fois théoriques et pratiques. D'une part nous développons de nouvelles méthodes d'ordonnancement et d'allocation en algèbre linéaire hétérogène et en télécommunications, de nouvelles techniques de distribution des données pour des bibliothèques de calcul distribué, et de nouvelles stratégies de transformation de codes (fusion de boucles et pipeline logiciel). D'autre part, nous développons également des outils de programmation parallèle, à travers les prototypes logiciels Scilab, TransTool, Nestor et SSCRAP.*

6.1.1 Algorithmique sur réseau hétérogène de processeurs

Pour distribuer les tableaux d'une application à paralléliser sur un réseau hétérogène de stations de travail, la solution bloc-cyclique traditionnelle n'est pas adaptée. Pour prendre en compte les différentes vitesses des processeurs, on peut penser à une distribution dynamique des données. Mais les stratégies dynamiques peuvent conduire à de mauvais résultats pour deux raisons : (i) le coût des communications liées à d'éventuelles redistributions rendues nécessaires par l'allocation des tâches ; (ii) l'inactivité forcée des processeurs à cause des contraintes de dépendances. Soulignons que la deuxième contrainte (ii) devient critique pour des plates-formes de calcul hétérogène : comme nous l'avons montré, les stratégies dynamiques gloutonnes habituellement employées sont vouées à l'échec dans ce contexte parce que les calculs des processeurs ne soient pas indépendants : très vite, l'exécution ne progresse qu'au rythme du processeur le plus lent.

Nous proposons une approche statique par phases de calcul, en modifiant l'allocation après chacune des phases pour tenir compte d'éventuels changements de vitesse des processeurs. Au sein de chaque phase, nous optimisons la charge des processeurs en fonction de leurs vitesses (ou des estimations courantes de celles-ci), et nous minimisons les communications.

Ces travaux initiés en 1999 ont connu d'importants développements au cours de l'année 2000.

Résultats de complexité Notre motivation initiale est l'étude du produit de deux matrices sur des plates-formes hétérogènes, soit de simples réseaux de stations de vitesses différentes,

soit (de manière plus ambitieuse) des grappes de tels réseaux. Intuitivement, le problème est d'équilibrer la charge sur ces ressources tout en minimisant le volume des communications.

Le problème abstrait peut se formuler ainsi : comment découper le carré unité en p rectangles d'aires données s_1, s_2, \dots, s_p (telles que $\sum_{i=1}^p s_i = 1$), de manière à minimiser (i) soit la somme des périmètres des p rectangles (ii) soit le plus grand périmètre de ces p rectangles ? Pour comprendre cette formulation abstraite :

- Les rectangles correspondent aux données allouées à chacun des p processeurs et l'aire s_i du i -ème rectangle correspond à la vitesse (normalisée) du i -ème processeur. Le fait que les aires soient fixées ainsi garantit l'équilibrage de charge.
- Chaque processeur échange des données horizontalement et verticalement. La taille de ces données est égale au demi-périmètre du rectangle alloué au processeur.
- Le volume total de toutes les communications est donc proportionnel à la somme des périmètres des rectangles, ou encore à la longueur des traits pour dessiner ces rectangles (jolie interprétation graphique!).
- Si on suppose les communications séquentielles (comme sur un réseau Ethernet par exemple), on cherche alors à minimiser la première fonction objective (i). Par contre si on suppose que certaines communications peuvent avoir lieu en parallèle (comme sur un réseau rapide dédié), on peut chercher à minimiser la plus longue d'entre elles, d'où la fonction objective (ii). Bien sûr tous les compromis intermédiaires entre ces deux cas extrêmes mais importants sont possibles.

Par ailleurs, on peut vouloir restreindre le partitionnement en colonnes de rectangles, afin de faciliter d'éventuelles redistributions ultérieures. On obtient alors quatre problèmes, les deux problèmes originaux PERI-SUM et PERI-MAX, et leur restrictions aux partitionnements en colonnes COL-PERI-SUM et COL-PERI-MAX. Nous avons démontré que (les problèmes de décision associés à) trois de ces problèmes étaient NP-complets, tandis que le dernier, COL-PERI-MAX, était polynômial. Pour COL-PERI-MAX, nous avons donné un algorithme de calcul de la solution optimale, et pour les trois autres nous avons proposé diverses heuristiques avec des bornes de garantie. Il s'agit là des premiers résultats de complexité du domaine : la littérature était riche en approches heuristiques de toute sorte, mais sans quantification théorique de leur performance (sans même de justification théorique, en fait, puisque les problèmes auraient pu être tous polynômiaux!).

Les preuves de tous ces résultats sont longues et techniques. Mais plus que les preuves elles-mêmes, c'est le résultat qui est intéressant : nous avons pu montrer la complexité inhérente des problèmes d'algorithmique hétérogène en algèbre linéaire.

ScaLAPACK sur grappe hétérogène Nous sommes désormais en mesure de proposer une stratégie d'allocation efficace pour les noyaux d'algèbre linéaire dense, dans le cas où la machine cible est un réseau hétérogène de stations de travail. En effet, nous avons pu étendre les techniques heuristiques employées pour le produit de matrices au cas des algorithmes de factorisation. La configuration virtuelle en colonnes de processeurs est beaucoup plus

extensible que la solution mono-dimensionnelle que nous avons proposée précédemment. Plusieurs expérimentations sous MPI sont venues corroborer les prévisions théoriques de manière très satisfaisante.

Pour l'approche semi-statique avec redistributions périodiques, nous avons obtenu plusieurs résultats théoriques, dont l'optimalité d'un algorithme glouton d'échanges partiels locaux. La réalisation d'un petit prototype logiciel est en cours.

Grille bi-dimensionnelle hétérogène Nous avons terminé l'étude amorcée l'an passé en démontrant la NP-complétude de l'équilibrage de charge sur une grille bi-dimensionnelle, NP-complétude qui nous avait résisté jusque là.

Algorithmique du métacomputing Nos travaux actuels s'orientent vers le problème de l'allocation des données en algèbre linéaire dans le cas de configurations matérielles plus complexes, composées d'une collection distribuée de grappes hétérogènes. Nous avons proposé une première modélisation pour ce type d'architectures et nous étudions actuellement plusieurs techniques d'allocation hiérarchique.

Conclusion L'algorithmique hétérogène s'est révélée à la fois difficile et intéressante, et comble notre vie de chercheur. Nos travaux ont conduit à plusieurs publications, et d'autres sont en cours de rédaction et soumission.

6.1.2 Algorithmique des télécommunications

Cette activité nouvelle dans *ReMaP* est appelée à se développer fortement, grâce aux recrutements d'Isabelle Guérin Lassous et de Nicolas Schabanel.

Le modèle CGM Le modèle CGM (Coarse Grained Multicomputer) est un modèle parallèle à gros grain récent qui prend en compte certaines caractéristiques réelles des machines parallèles actuelles tout en englobant le plus de plates-formes possibles. Nos travaux ont porté sur une étude expérimentale concernant les algorithmes parallèles de traitement de graphes et le modèle CGM. Nous nous sommes donc intéressés à la conception d'algorithmes parallèles de traitement de graphes ainsi qu'à leur implantation sur des architectures diverses.

Cette année, nous avons proposé plusieurs algorithmes pour le problème du list ranking présentant de bonnes complexités théoriques. C'est un problème de base en algorithmique parallèle des graphes puisqu'il est utilisé comme sous-routine dans la plupart des algorithmes parallèles sur les graphes. Nous avons proposé un algorithme de coloriage de graphe dans ce modèle. Nous avons aussi proposé des algorithmes de permutation aléatoire d'entiers qui respectent les contraintes du modèle CGM. Ces algorithmes permettent de réaliser une redistribution aléatoire des données comme il est souvent demandé dans les algorithmes probabilistes.

Concernant la partie expérimentation, un certain nombre d'algorithmes ont été implantés sur une machine parallèle T3E et deux grappes de PCs (une reliée par un réseau Fast Ethernet

et l'autre par un réseau Myrinet). Les algorithmes implantés concernent les algorithmes proposés cette année, ainsi que d'autres algorithmes soit conçus les années précédentes (somme préfixe, graphes de permutation, graphes d'intervalles), soit déjà existants dans la littérature. Toutes ces implantations ont fait l'objet d'une bibliothèque appelée SSCRAP qui se situe sur deux niveaux : le premier implante les fonctionnalités du modèle CGM et le deuxième concerne les algorithmes CGM de traitement de graphes (le deuxième niveau faisant appel au premier). SSCRAP est basée sur la bibliothèque MPI. Ce travail a montré que ce modèle permet de faire du calcul efficace sur les réseaux locaux homogènes à haut débit et les machines parallèles car peu de communications sont réalisées et au total peu de données sont échangées sur le réseau.

Nous allons poursuivre nos travaux en cours sur le coloriage et l'étendre au problème d'allocation de fréquences dans les réseaux sans fil. Ce problème peut être résolu par des techniques de coloriage de graphe. Nous espérons pouvoir proposer des solutions à des réseaux ayant des configurations très complexes (proximité d'un pays étranger, présence de montagnes et de la mer) et qui requièrent donc beaucoup de calculs et de stockage mémoire. Ce travail a fait l'objet d'une proposition de projet au programme Aurora (échange France-Norvège).

D'autre part, les résultats ont montré que le médium de communication était peu sollicité avec ce modèle. Cette approche semble donc prometteuse et bénéfique pour certaines applications qui demandent d'économiser de la bande passante. C'est le cas des réseaux radio qui ont un médium de communication partagée et leur bande passante est réduite. Par conséquent, les protocoles qui limitent l'utilisation du réseau de communication ont de grandes chances d'être efficaces. Nous étudions actuellement dans quelle mesure ce modèle peut être utilisé pour la proposition de protocoles efficaces dans les réseaux sans fil.

Planification des réseaux locaux sans fil Dans le cadre de l'Action concertée incitative *Jeunes chercheurs* 2000, nous avons commencé à travailler sur la modélisation et l'optimisation des réseaux locaux sans fil. Le but est de proposer une modélisation spatiale en intérieur (*indoor*) suffisamment réaliste qui permette d'une part la planification de tels réseaux et qui permette d'autre part d'évaluer les performances des protocoles de réseaux (comme l'accès au médium radio, le routage, etc). Cette évaluation devrait nous permettre soit d'affiner les protocoles en place, soit d'en proposer de nouveaux.

Ce travail qui vient de commencer propose déjà une modélisation de la propagation hertzienne en intérieur en deux dimensions à l'aide d'une discrétisation de l'espace de propagation en cellules. Le passage en trois dimensions se fait naturellement avec une forte augmentation des calculs. Le travail porte actuellement sur la diminution de ces calculs.

Routage dans les réseaux L'étude du routage dans les réseaux se fait dans deux directions. La première concerne le projet européen ProxiTV auquel nous participons. Ce projet a pour objectif de mettre en place une infrastructure proxy pour la télévision interactive sur Internet. L'idée est d'utiliser les réseaux haut débit (type câble, ADSL, etc) qui relieront les utilisateurs et les serveurs proxy. Les données proposées par les fournisseurs de contenu (WebfreeTV, Eurosport) sont initialement localisées sur des portails. Puis de manière régulière (à défi-

nir par les partenaires du projet) ces données sont routées sur les différents serveurs proxy. L'utilisateur accédera alors aux données via le réseau haut débit auquel il est rattaché et récupérera donc très rapidement les contenus qu'il désire. Notre travail au sein de ce projet est de proposer des algorithmes de routage des portails aux serveurs proxy, c'est-à-dire dans des réseaux hautement hétérogènes. Le premier travail sera de modéliser ce type de réseau, puis de proposer des algorithmes de routage qui optimisent une (ou des) fonction(s) de coût (ces fonctions sont à définir par les partenaires ; elles peuvent être le délai le plus petit, le tarif le moins cher, etc). Un démonstrateur doit être réalisé et un test grandeur réel sur un des réseaux d'Austria Telekom doit être effectué.

La deuxième direction concerne le routage des communications multi-points dans les réseaux ad hoc. Un réseau ad hoc est un réseau sans fil dans lequel il n'y a aucune infrastructure fixe, les ordinateurs mobiles pouvant être des routeurs potentiels. Ces réseaux ont fait l'objet d'un groupe de travail à l'IETF appelé Manet. Des protocoles de routage ont été proposés et sont en cours de normalisation. Nous essayons de voir actuellement dans quelle mesure nous pouvons utiliser les techniques de protocoles de routage point-à-point pour les protocoles de routage multi-points.

Qualité de service Cette année, nous nous sommes intéressés au problème de réservation de bande passante pour les réseaux ad-hoc. Ce travail, récent, a permis notamment de montrer qu'une modélisation de la réutilisation spatiale permettait d'exprimer le problème de réservation de bande passante en un problème d'optimisation original (le problème du sac-à-dos multi-dimensionnel). Nous avons montré que les simplifications apportées par le cadre des réseaux radio laissent le problème de réservation de bande passante NP-complet. Nous avons proposé de premières heuristiques et calculé les bornes associées concernant la taille de la bande passante utilisée.

Nous continuons à travailler activement sur cette problématique : nous espérons proposer de nouvelles heuristiques ayant de meilleures approximations et nous comptons développer un simulateur afin de vérifier expérimentalement la validité de nos analyses théoriques.

Dissémination de données Nous avons poursuivi nos travaux sur la dissémination de données (un problème de minimisation issu d'un protocole de diffusion de données sur les réseaux sans fil type satellite, téléphonie mobile, systèmes embarqués, etc.). Notre collaboration avec Neal Young (Dartmouth College, New Hampshire, USA, 2 séjours d'un total de 3 mois 1/2) a porté ces fruits et nous avons pu construire le premier schéma d'approximation pour ce problème. Nous avons également exploré un modèle préemptif pour la dissémination de données. Ces résultats sont rassemblés dans la thèse de N. Schabanel intitulée *Algorithmes d'approximation pour les télécommunications sans fil : dissémination de données et allocation statique de fréquences*.

Nous allons poursuivre nos travaux au cours de plusieurs missions et collaborations dans le cadre du DIMACS *Special Focus on Next Generation Networks Technologies and Applications* à DIMACS (Rutgers University, New Jersey) et ATT Labs (Florham Park, New Jersey).

6.1.3 Transformation de programmes

À l'heure actuelle, le parallélisme est exploité à tous les niveaux, bien sûr dans les machines clairement identifiées comme parallèles, mais également au sein même des microprocesseurs, microprocesseurs généraux superscalaires, VLIW (Very Long Instruction Word) et également processeurs dédiés. Pour toutes ces plates-formes, aussi différentes soient-elles, l'étude des transformations de programmes révélant du parallélisme et leur automatisation est notre centre principal d'intérêt. Nos dernières études ont principalement abordé deux transformations : la distribution de boucles (et son inverse la fusion) en présence de *dépendances de contrôle* et le *décalage d'instructions*. Cette dernière transformation a de nombreuses applications, à la fois en synthèse d'architectures et en parallélisation automatique : nos principaux résultats la concernant portent sur la détection de boucles parallèles et le pipeline logiciel.

Distribution de boucles et dépendances de contrôle Nous avons étudié la parallélisation de boucles dans des programmes non statiques, c'est-à-dire possédant un flot de contrôle pouvant varier à l'exécution en fonction des données (typiquement des codes avec des IF). Notre étude a porté en outre sur une extension de l'algorithme d'Allen, Callahan et Kennedy, pour laquelle nous pouvons contrôler la dimension des tableaux à introduire dans le code final pour garantir le bon calcul des bornes de boucles et des conditionnelles. Cet algorithme a été présenté à la conférence Euro-Par 2000.

Décalage d'instructions et boucles parallèles Nous avons cherché à déterminer les possibilités de parallélisation de boucles à l'aide de simples décalages. Cette approche permet de trouver des ordonnancements que les techniques classiques ne trouvent pas (en tirant parti des dépendances indépendantes de la boucle, en d'autres termes de séquentialité interne au corps de la boucle) et conduit à une transformation simple et efficace du programme, ce qui n'est pas toujours le cas lors de transformations plus complexes (par exemple, les transformations affines plus générales posent des problèmes d'élimination de gardes). Nous avons établi la NP-complétude du problème (résultat en contradiction – sauf si $P=NP$ – avec un algorithme (erroné) polynômial présenté dans la littérature) et avons proposé une formulation exacte de celui-ci sous forme d'un système de contraintes linéaires entières.

Décalage d'instructions et pipeline logiciel Notre étude concernant le pipeline logiciel a porté sur l'approche dite décomposée qui consiste à considérer un ordonnancement cyclique (effectuant le pipeline logiciel) comme la combinaison d'un décalage d'instructions (ou retiming) et d'un ordonnancement acyclique (compaction de boucles). L'intérêt de ce découpage en deux phases est notamment de pouvoir tirer parti des résultats existants tant du côté des techniques d'ordonnancement cyclique que des techniques de décalage proches algorithmiquement des optimisations de flots dans les graphes.

D'un point de vue théorique, nous avons proposé un algorithme polynômial permettant de minimiser le nombre total de contraintes restant lors de la phase de compaction et qui peut très simplement se combiner à la minimisation du chemin critique ou à une contrainte plus faible par ajout d'arcs dans le graphe de dépendances. D'un point de vue pratique, nous

avons confronté notre technique à des tests d'efficacité sur un large ensemble de graphes de dépendance, à l'aide d'un petit logiciel que nous avons développé (voir la section 5.1). L'importance de ces tests était justifiée d'une part par l'absence totale de résultats expérimentaux sur les performances réelles d'une approche à base de décalage, d'autre part par la nécessité d'étudier l'influence du décalage sur le choix d'une technique de compaction particulière. Ces tests ont permis de révéler les points forts et les problèmes restants dans notre approche. Nous avons pu constater une bonne efficacité absolue de la technique qui, de plus, s'avère excellente dans les cas de pénurie de ressources. En revanche, des problèmes dans le choix du décalage sont apparus pour les ordonnancements de période très courte (inférieure à la durée d'une instruction) et la technique n'est pas encore à même de limiter l'augmentation du nombre de registres utilisés. Alors que le premier problème peut être contourné relativement facilement, le deuxième s'avère plus complexe et fait l'objet de nos études courantes. Tous ces résultats viennent d'être publiés dans IJPP (*International Journal of Parallel Programming*).

6.1.4 Serveurs de calculs et optimisation d'algorithmes numériques parallèles

Serveurs de calculs Le but de l'ARC OURAGAN est de développer une version parallèle du logiciel *Scilab*. Notre but est de conserver l'interactivité de cet outil tout en améliorant les performances et en cachant le plus possible l'emploi du parallélisme à l'utilisateur. Il s'agit de permettre au programmeur d'avoir différents niveaux d'utilisation du parallélisme : soit le parallélisme est totalement caché grâce à une surcharge des opérateurs d'algèbre linéaire, soit le programmeur a un contrôle sur la distribution des données et des calculs. Deux approches ont été retenues : soit nous dupliquons des processus Scilab sur les divers processeurs mis à notre disposition, soit nous utilisons des serveurs de calculs disposés sur les grappes accessibles via le net.

Dans le premier modèle, après avoir intégré la bibliothèque de passage de message PVM, nous avons travaillé sur l'interfaçage de bibliothèques parallèles comme ScaLAPACK. Il s'agit d'ajouter à Scilab des types de données distribuées et de surcharger les opérateurs arithmétiques standards par des opérations matricielles utilisant des bibliothèques parallèles. Cette approche a plusieurs avantages puisque qu'elle permet un développement assez rapide grâce à l'intégration des bibliothèques directement dans l'outil mais par contre, elle nécessite des modifications des types si nous voulons ajouter d'autres bibliothèques.

Dans le second modèle, nous utilisons une approche client-agent-serveur. Un client souhaitant effectuer un calcul demande à un agent quel est le serveur le plus adapté pour effectuer ce calcul (à la fois en terme de capacité logicielle et de puissance de calcul ou de capacité mémoire). Après avoir consulté les serveurs dont il a la responsabilité, l'agent décide (grâce à une heuristique d'ordonnement) quel est le serveur qui doit effectuer le calcul et le signale au client. Le client envoie alors ses données au serveur qui effectue le calcul et retourne le résultat. Nous sommes partis du logiciel Netsolve développé à l'université du Tennessee. Netsolve est un logiciel qui fonctionne déjà selon ce modèle mais qui demande à être amélioré. Dans un premier temps, nous avons ajouté à Netsolve la persistance de données. Plutôt que de ramener les résultats à chaque fois sur le client, nous laissons les données sur place et nous les réutilisons dans les calculs futurs. Cela permet d'améliorer les performances en réduisant la latence. Ensuite nous avons travaillé sur deux outils annexes : SLiM et FAST. SLiM

(*Scientific Libraries Metaserver*) est une base de données permettant à l'agent de trouver sur les serveurs la bibliothèque qui permettra de résoudre un problème. FAST (*Fast Agent System Timer*) est un outil qui permet de dire à l'agent quel est le serveur le plus adapté en terme de performances et de capacité mémoire. Utilisant le logiciel NWS (*Network Weather Service*), FAST peut calculer le coût de migration des données entre le client et les serveurs ou entre les clients entre eux. Il utilise également des benchmarks exécutés au démarrage du serveur qui permettent d'avoir un modèle des coûts de calcul pour les diverses routines disponibles.

Ces travaux sont effectués en collaboration avec les projets Métalau (INRIA Rocquencourt), Résédas (INRIA Lorraine) et les laboratoires LaBRI (Bordeaux), LaRIA (Amiens) et LIFC (Besançon).

Optimisation d'algorithmes numériques parallèles Nous avons poursuivi nos travaux sur l'utilisation des recouvrements calculs/communications dans les algorithmes parallèles. Nos travaux précédents concernaient des algorithmes réguliers utilisant des schémas pipelines relativement simple (pipelines mono-dimensionnels). Ensuite, nous nous sommes intéressés à des schémas pipelines simples (toujours mono-dimensionnels) sur des structures de données irrégulières (algorithme du Shear-Warp). Nous nous intéressons maintenant à des schémas pipelines multi-dimensionnels sur des structures de données régulières. La difficulté vient du fait qu'un recouvrement existe déjà entre les calculs du fait de la distribution multi-dimensionnelle. Le gain n'est plus aussi important qu'avec les pipelines mono-dimensionnels. Nous avons commencé par évaluer le gain issu de tels algorithmes. Dans ce but, nous avons étudié une application benchmark du programme américain ASCI, le Sweep3D. Cette application effectue des calculs par vagues sur des matrices à trois dimensions. Il existe de nombreuses manières de pipeliner les calculs et les communications et de distribuer les données. On peut utiliser une distribution mono, bi ou tridimensionnelle et plusieurs dimensions peuvent être pipelinées.

Nous avons développé des algorithmes pour le découpage de graphes de tâches à l'aide de techniques macro-pipelines. Nous supposons que l'algorithme parallèle est constitué d'une séquence d'appels à des routines d'algèbre linéaire comme le produit de matrices ou la résolution de systèmes triangulaires. Nous supposons également qu'il existe des dépendances entre les routines du fait des matrices utilisées entre elles. Nous plaçons chaque routine sur un processeur différent. Ces routines peuvent être découpées de diverses manières suivant les dimensions des matrices utilisées dans les calculs. Ces découpages vont modifier les schémas de communication et conduire à un pipeline des calculs et des communications. Ce pipeline peut conduire, s'il respecte les grains de calcul et de communications adéquats, à des gains de performance importants. Nous avons proposé un algorithme de découpage d'un tel graphe pipelinant calculs et communications.

Les algorithmes parallèles peuvent être séparés en deux catégories : les algorithmes utilisant un paradigme de programmation *data-parallèle* et ceux utilisant le *parallélisme de tâches*. Dans le premier modèle, les données utiles au calcul sont découpées selon la grille virtuelle de processeurs et ceux-ci effectuent les mêmes opérations sur des données différentes ; il s'agit donc d'une approche SPMD. Dans le second modèle, le programme parallèle est découpé en tâches et ces tâches sont réparties sur les processeurs selon les dépendances qui les

lient. Nous travaillons maintenant sur l'algorithmique parallèle mixte. Dans ce modèle, nos programmes sont constitués de tâches qui sont elles-mêmes data-parallèles. Notre premier algorithme cible est le produit de matrices utilisant les méthodes de Strassen ou Winograd. Les parallélisations précédentes utilisaient généralement soit le data-parallélisme, soit le parallélisme de tâches. Nous avons prouvé que nous pouvions avoir un gain en performances grâce au parallélisme mixte en optimisant le placement des données et les redistributions entre les sous-grilles de processeurs.

6.2 Environnements d'exécution parallèles et distribués

Participants : Gabriel Antoniu, Olivier Aumage, Alice Bonhomme, Luc Bougé, Vincent Danjean, Alexandre Denis, Jean-François Méhaut, Guillaume Mercier, Jean-Christophe Mignot, Raymond Namyst, Loïc Prylli.

Résumé :

Les activités dans ce thème ont été cette année consacrées à l'évolution de l'environnement d'exécution PM2 d'une part, et à l'évolution de certains aspects du système BIP d'autre part.

L'activité autour de PM2 s'est déroulée suivant trois directions principales : l'amélioration de la structure de base (meilleur support des réseaux haute performance, meilleure intégration threads/communications), l'extension des fonctionnalités (mémoire virtuelle partagée) pour l'exécution distribuée de byte-codes Java et l'évolution vers les grappes hétérogènes (gestion des communications multi-protocoles). L'ensemble de ces travaux a été répercuté sur la suite logicielle PM2 (PM2, Madeleine, Marcel et DSM-PM2).

En ce qui concerne BIP, de nouveaux mécanismes de recouvrement calcul/communication ont été étudiés dans MPI-BIP et MPI-GM. D'autres améliorations au niveau du module noyau de BIP permette d'améliorer sa sécurité et sa fonctionnalité.

Enfin au niveau applicatif, des réalisations ont été effectuées dans le domaine des serveurs multimédia.

6.2.1 Environnement PM2

Participants : Luc Bougé, Vincent Danjean, Jean-François Méhaut, Raymond Namyst.

Cette année marque à la fois le renforcement de l'environnement PM2 en tant que support exécutif pour les grappes de stations et son ouverture sur de nouvelles architectures ainsi que de nouvelles applications. En 1999, nous avons dégagé une architecture intéressante s'appuyant d'une part sur une couche de communication générique (*Madeleine II*) optimisée pour les réseaux rapides et d'autre part sur une bibliothèque de threads mixte (Marcel) capable d'exploiter les architectures SMP de manière très efficace. Sur cette base, nous avons travaillé cette année à l'amélioration de divers aspects concernant l'intégration des threads et des communications.

Activations. Nous avons proposé une extension du modèle des *Scheduler Activations* dans le contexte particulier des applications de calcul sur machines SMP (travaux de Vincent

Danjan). L'idée du modèle original consiste à étendre le noyau du système afin d'offrir à un ordonnanceur de threads de niveau utilisateur un support permettant la gestion efficace des appels systèmes bloquants. Ce modèle souffrait cependant de multiples limitations, d'un manque de généricité et surtout d'une mise en œuvre intrinsèquement peu efficace. En privilégiant un type d'exploitation *mono-application* de la machine, nous avons proposé et implanté une nouvelle version de ce mécanisme capable de traiter la totalité des appels système tout en optimisant fortement la gestion et le nombre d'interactions depuis le noyau vers l'espace utilisateur (ce type d'interaction, symétrique d'un appel système classiques, est appelé *upcall*). L'implantation a été effectuée dans le noyau Linux 2.2.x et la bibliothèque multithread Marcel. Sur une telle plate-forme, l'exécution des applications PM2 peut s'effectuer sans recourir aux classiques opérations de scrutation pour la plupart des protocoles réseaux.

Scrutations dirigées par l'ordonnanceur. Nous avons proposé une interface unificatrice pour réaliser l'intégration fine des entrées/sorties (en particulier les opérations liées aux communications) avec l'ordonnanceur de threads Marcel. Dans le cas de primitives d'entrées/sorties bloquantes (utilisant les interruptions), le mécanisme d'activations décrit précédemment est automatiquement utilisé sous Linux, ou alors la délégation de l'opération à un thread noyau est effectuée pour les autres systèmes. Dans le cas contraire, c'est-à-dire si les opérations requièrent une scrutation active de la part du processeur, alors l'ordonnanceur de threads est sollicité pour prendre en charge les opérations de scrutation et pour éventuellement factoriser les opérations de même nature. Ceci permet non seulement d'assurer une période de scrutation garantie, mais aussi d'éliminer de nombreux changements de contextes inutiles lorsqu'une scrutation réseau est effectuée. Il en résulte une bien meilleure réactivité des applications PM2 qui peuvent ainsi traiter les requêtes réseaux au plus tôt.

6.2.2 Communications multi-protocoles

Participants : Olivier Aumage, Luc Bougé, Alexandre Denis, Jean-François Méhaut, Guillaume Mercier, Raymond Namyst.

Madeleine sur architectures hétérogènes. Notre travail s'est articulé autour de l'évolution de la bibliothèque de communication *Madeleine II* vers le support efficace des architectures de type *grappes de grappes* (travaux d'Olivier Aumage). Cette étude a nécessité le développement d'outils auxiliaires spécialisés (chargeur, contrôleur de session, routeur) afin de maîtriser la complexité de la mise en œuvre de sessions de calcul distribuées sur de multiples grappes de stations. Ces travaux ont été menés conjointement à la phase de finalisation et de diffusion de la bibliothèque *Madeleine II* dans son orientation initiale, à savoir le support de communications multi-protocoles sur des grappes homogènes de stations de travail. Dans ce cadre, *Madeleine II* a fait l'objet d'une refonte partielle de son architecture interne et a intégré quatre nouveaux pilotes pour des interfaces réseaux aussi diverses que SISI/SCI, BIP/Myrinet, SBP/Fast-Ethernet et MPI.

Routage au sein des grappes de grappes. L'exploitation efficace d'architectures hétérogène de type *grappes de grappes* exige que le sous-système de communication fournisse des fonctionnalités de routages intégrées. Le point crucial concerne bien évidemment les machines *passerelles* qui possèdent plusieurs cartes d'interfaces et qui doivent assurer rapidement le transfert des messages d'un type de réseau à un autre. Cela signifie soutenir un débit élevé et préserver une latence faible. Nous avons donc conçu un mécanisme permettant à Madeleine de réaliser un routage élémentaire très efficace de manière interne, c'est-à-dire sans que les données ne remontent dans les couches logicielles supérieures (travaux d'Olivier Aumage). Outre son efficacité (les débits observés montrent une quasi-saturation du bus PCI sur des grappes de PC), ce mécanisme apporte aussi une certaine souplesse d'utilisation puisque des threads applicatifs peuvent également s'exécuter sur les machines passerelles.

Intégration de Madeleine dans d'autres bibliothèques de communication. La bibliothèque Madeleine, bien que proposant une interface générique au programmeur, possède des caractéristiques remarquables en terme d'efficacité avec notamment un temps de transfert minimal de $4 \mu s$ sur un réseau SCI (le réseau lui-même ayant une latence d'environ $2 \mu s$). Elle a donc logiquement été intégrée comme plate-forme de communication dans plusieurs logiciels : bien évidemment dans PM2, le support d'exécution multithread distribué développé au sein de l'équipe, mais aussi dans les bibliothèques de communication MPICH (travaux de Guillaume Mercier) et Nexus (travaux d'Alexandre Denis) pour lesquelles elle fournit un service de communication multiprotocole efficace. MPICH est une bibliothèque de communication très largement utilisée dans le domaine du calcul scientifique distribué. Notre implémentation de MPICH au-dessus de *Madeleine II* donne les meilleurs résultats actuels sur le réseau SCI. Quant à Nexus, il s'agit de la bibliothèque de communication utilisée par le projet Globus initié à l'Argonne Nat. Lab. et à l'Université de Californie (USA) et dont le but est de fournir une grille de calcul, ou *metacomputing grid*, à échelle mondiale.

6.2.3 Évolution de PM2 pour le support du langage Java

Participants : Gabriel Antoniu, Luc Bougé, Raymond Namyst.

Système DSM-PM2. Nos discussions avec Assaf Schuster (Technion, Haïfa) et Frank Mueller (Humboldt Univ., Berlin) nous ont conduits à envisager d'utiliser PM2 et son iso-allocateur comme outils de base pour un système de gestion mémoire virtuellement partagée (*Distributed Shared Memory, DSM*) spécialisé pour le multi-threading haute performance et la migration dynamique de threads. Nous avons donc conçu et implanté une couche *générique* de mémoire virtuellement partagée multithread (DSM-PM2, travaux de Gabriel Antoniu), supportant plusieurs modèles de cohérence : la cohérence séquentielle, la cohérence relâchée ainsi que la cohérence Java. En outre, l'utilisateur peut définir ses propres modèles de cohérence et/ou implanter ses propres protocoles pour une cohérence fixée. Notons qu'en ce qui concerne la cohérence séquentielle, un protocole original basé sur la migration de threads est fonctionnel et permet

d'envisager des stratégies adaptatives reposant sur l'utilisation alternée de la migration des pages et de la migration des threads.

Intégration dans la plate-forme Hyperion. Le système DSM-PM2, et notamment son support du modèle de cohérence mémoire Java, constitue le socle de l'environnement d'exécution Java distribué *Hyperion* développé par Phil Hatcher (Univ. New Hampshire, Durham, USA). Un premier prototype est déjà opérationnel. L'objectif est comme toujours de maintenir la portabilité des performances sur la plus grande gamme d'architecture de nœuds, de protocoles de communication et de réseaux d'interconnexion possible. Ce travail constitue le cœur de la thèse de Gabriel Antoniu. L'intégration des deux systèmes est actuellement en phase d'optimisation.

6.2.4 MPI au-dessus de PM2

Participants : Guillaume Mercier, Loïc Prylli, Raymond Namyst.

MPICH-Madeleine. est une version de MPICH utilisant *Madeleine II* comme support pour les communications. L'intérêt de la démarche réside dans l'exploitation par MPICH des capacités multiprotocoles de *Madeleine*. Le prototype implémenté est capable de passer dynamiquement d'un réseau d'interconnexion à un autre. Les protocoles supportés actuellement sont : TCP au-dessus de Fast-Ethernet, BIP au-dessus de Myrinet et enfin SISI au-dessus du réseau SCI. Les performances obtenues au-dessus de ces trois types de matériel sont très bonnes ; les comparaisons effectuées en termes de débit avec d'autres implémentations (monoprotocoles) de MPI au-dessus de réseaux similaires montrent que notre version de MPICH multiprotocole atteint des niveaux de performances équivalents et/ou supérieurs à ceux obtenus par ces autres implémentations.

6.2.5 Évolutions de BIP

Participant : Loïc Prylli.

Recouvrement calcul/communications. BIP (et GM) ont fait l'objet de travaux permettant d'implémenter sur l'hôte des protocoles qui vont s'exécuter en *tâche de fond* pendant que le processeur exécute des calculs non liés au réseau. L'objectif est de concilier le besoin de réactivité de ces protocoles (et donc d'utiliser des interruptions pour donner la main au code réseau), tout en préservant les faibles latences de BIP que l'on obtient par scrutation. Cela nous a conduit à utiliser une heuristique simple permettant d'éviter de générer des interruptions lorsque le code réseau a suffisamment régulièrement la main.

Gestion de la mémoire virtuelle. Les mécanismes d'enregistrement mémoire utilisés pour les transmissions à zéro-copie ont été intégrés de manière plus fine avec la gestion de la mémoire virtuelle dans Linux. Ceci a permis de les rendre compatibles avec la primitive `fork()` qui permet de créer de nouveaux processus dans un processus BIP.

6.2.6 Activités multimédia

Participants : Alice Bonhomme, Jean-François Méhaut, Jean-Christophe Mignot, Loïc Prylli.

Cache Web parallèle. Dans le cadre du projet *CHARM*, un cache parallèle extensible basé sur une grappe de PC a été réalisé. Ce cache repose sur l'utilisation d'un premier aiguilleur IP réalisé en utilisant les fonctionnalités *VirtualServer* de Linux. Cet aiguilleur distribue les connexions TCP vers un ensemble de *routeurs HTTP* permettant de distribuer les requêtes entre plusieurs caches séquentiels, suivant une fonction de partitionnement des URL.

Système de stockage pour la vidéo. Dans le cadre d'une bourse CIFRE (Alice Bonhomme) avec la société Sycomore-Aérospatiale-Matra, un système de stockage distribué haute performance pour la vidéo a été réalisé. Le point fort de ce système est la combinaison entre la tolérance aux pannes, y compris l'absence de perturbation temporaire de la distribution des flux en cas de panne, l'extensibilité (aucune opération centralisée), et un système de redondance basé sur la parité plus économique que la duplication. Ce système est d'ores et déjà utilisé pour des démonstrations industrielles de produits destinés à la commercialisation.

7 Contrats industriels (nationaux, européens et internationaux)

7.1 LHPC

Le *LHPC* (Laboratoire pour les Hautes Performances en Calcul) est un laboratoire commun de recherche sur les ordinateurs massivement parallèles et le calcul à haute performance créé entre l'INRIA (projet *ReMaP*), l'ENS Lyon, le CNRS et la société MATRA Systèmes & Information (*MS&I*). Il s'appuie sur le contrat de plan état-région avec la Région Rhône-Alpes. Le *LHPC* a été officiellement inauguré le 2 décembre 1996. La convention de collaboration a été renouvelée au 1^{er} janvier 1999, pour une période de quatre ans.

La vocation du *LHPC* est d'être un centre de compétence dans le domaine du calcul parallèle et de réunir des partenaires d'origines variées autour d'un projet académique et industriel aux larges ambitions. Les objectifs du *LHPC* peuvent se résumer selon 4 axes principaux.

- Participer à la réalisation d'une plate-forme de calcul modulaire et extensible.
- Concevoir et développer un environnement de programmation portable mais permettant l'exploitation optimale de la plate-forme d'expérimentation.
- Mettre la plate-forme d'expérimentation à la disposition des chercheurs et des industriels et leur offrir des services d'ingénierie pour le portage d'applications existantes ou le développement de nouvelles applications, ainsi que pour les activités de recherche associées.
- Assurer la formation continue des étudiants, chercheurs et ingénieurs.

Les plates-formes de calcul les plus récentes du LHPC sont des piles de PC : pile de Pentium MMX + Myrinet (12 processeurs) ; pile de Power PC + Myrinet (16 processeurs) ; pile de Pentium II + SCI (16 processeurs) ; pile de Pentium II + Myrinet (8 processeurs) ; pile de Alpha + Myrinet (4 processeurs) ; serveur multimédia (16 processeurs, 400 Giga-octets de disque). Ces plates-formes vont être reliées entre elles par des liens rapides pour former des collections de grappes (*grappes de grappes*). L'objectif est de permettre des expérimentations au niveau algorithmique et au niveau système dans ce nouveau contexte.

Les principales collaborations au sein du LHPC en 2000 ont été :

- le projet CHARM,
- le montage du projet SPIHD.

Les projets CHARM et SPIHD sont résumés ci-dessous.

Projet CHARM Le projet CHARM (*Cache haut débit pour les autoroutes multimédia*) est financé par le SERICS, programme *Autoroutes de l'information*. Il est présenté par un consortium dirigé par le Centre Érasme, un établissement de recherche public soutenu par le Conseil général du Rhône et l'Éducation nationale pour étudier l'utilisation des nouvelles technologies multimédia sur réseaux à haut débit à des fins pédagogiques. Le LIP et *MS&I* sont associés dans ce projet pour le développement d'un système de caches Web haute performance pour des boucles locales haut débit. J.-F. Méhaut, J.-C. Mignot et L. Prylli assurent la réalisation d'une maquette de cache hiérarchique parallèle faiblement couplée et son optimisation. L. Bougé assure la responsabilité scientifique du projet en ce qui concerne le LIP. Alice Bonhomme (encadrée par L. Prylli) et Frédérique Chaussumier (encadrée par F. Desprez) entament leurs deuxièmes années de thèse sur des bourses Cifre LHPC dans ce cadre. Le projet CHARM s'est terminé en octobre 2000. La revue aura lieu en décembre 2000. Plusieurs des rapports de contrat rédigés par Jean-Christophe Mignot proposent des états de l'art sur certains aspects de ce domaine. Ils ont été diffusés sous forme de rapport INRIA et soumis pour publication.

Projet SPIHD Le projet SPIHD (*Services et Programmes pour l'Internet Haut Débit*) est mené en collaboration avec *MS&I*, des professionnels de l'audiovisuel (France 3, CanalWeb, la Cinquième) et des partenaires de la région lorraine (le LORIA et SEM-Câble de l'Est). L'objectif du projet est de développer une nouvelle approche de production et de diffusion de contenus multimédia et de journaux télévisés sur l'Internet haut débit et les boucles locales. Les expérimentations sont menées sur Nancy et sa région. Les tâches affectées à ReMaP concernent l'étude des nouvelles technologies de codage/décodage vidéo.

Une première étude a été réalisée sur les techniques de compression pour la vidéo et les techniques d'encodage et de décodage (*codecs*) les plus récentes disponibles. Les techniques de compression vidéo des normes JPEG, MPEG2 et MPEG4 y sont décrites en détail. Ce document est un rapport de contrat, il sera disponible en rapport de recherche du LIP après approbation des différents partenaires du projet SPIHD. D'autre part, une seconde étude a été menée sur les techniques d'indexation de vidéo et les principaux formats en émergence dans ce domaine. Cette étude a mis en évidence le langage XML qui présente l'avantage

d'être en passe de devenir le standard de référence. Un prototype de démonstration est en cours de réalisation montrant les possibilités de recherche d'informations d'index au format XML sur une base de données vidéo.

7.2 Collaboration avec Myricom

La collaboration avec Myricom s'est poursuivie pour la troisième année avec un nouveau séjour de L. Prylli dans la société. La prestation fournie à Myricom a portée sur les points suivants :

- amélioration des mécanismes d'enregistrement mémoire (compatibilité avec `fork()`),
- introduction dans GM et MPI-GM d'un mécanisme permettant de mélanger les mécanismes de scrutation et l'utilisation des interruptions pour éliminer les problèmes de progression des communications asynchrones lorsque le processeur est occupé dans des calculs,
- expertise système Linux, notamment support pour les nouvelles versions du noyau.

7.3 Contrat INRIA/Microsoft

ReMaP participe, avec trois autres projets INRIA (Apache, Reso, Sirac), à un programme de collaboration avec la société Microsoft. L'objectif de ce programme est de porter un ensemble de logiciels issus des quatre projets INRIA sur le système d'exploitation Windows-NT. En ce qui concerne ReMaP, le logiciel PM2 et ses composants internes (Marcel, Madeleine) ont été proposés à Microsoft. Ce travail technique de développement a été pris en charge par Patrice Dargenton (ingénieur-expert recruté par l'ENS Lyon). Le noyau de processus légers Marcel a ainsi été porté sur Windows NT et des comparaisons de performance avec le système Linux sont en cours. En ce qui concerne le noyau de communication Madeleine, les travaux débiteront dès que les grappes de stations sous Windows-NT seront installées et opérationnelles. Cette collaboration avec Microsoft a permis de renforcer les liens entre les projets INRIA impliqués.

7.4 Accord-cadre INRIA avec la société Alcatel

Un contrat de coopération est sur le point d'être signé (courant novembre 2000) entre l'entreprise Alcatel et l'INRIA (projet lyonnais ReMaP et projet rennais PARIS) portant sur l'étude et la conception de mécanismes de communication pour la réalisation de bus logiciels performants basés sur l'interface CORBA. L'objectif est d'élaborer une mise en œuvre de CORBA capable d'exploiter finement des technologies réseaux rapides telles que VIA, Myrinet ou encore SCI. Le projet ReMaP propose d'étudier l'intégration de la librairie de communication Madeleine dans des environnements CORBA. Plus précisément, il s'agit d'étudier le couplage de Madeleine avec l'interface de transport générique OCI (Open Communication Interface) dont l'intérêt principal réside dans la réutilisabilité de son implantation : on obtient une solution indépendante des ORB compatibles et ne nécessitant pas d'accès à leur

code source. Le cas échéant, des extensions de Madeleine seront effectuées pour faciliter l'intégration.

8 Actions régionales, nationales et internationales

8.1 Actions nationales

Un grand nombre d'actions nationales du projet ReMaP se font dans le cadre du GDR ARP (*Architecture, Réseaux et systèmes, Parallélisme*) dirigé par Michel Diaz (LAAS, Toulouse) avec l'aide de Luc Bougé et Daniel Litaize (IRIT, Toulouse).

ASP MENRT iHPerf'98. En juillet 1997, l'inter-GDR a présenté au MENRT une série de propositions d'Actions sur programme (ASP). Parmi celles-ci, L. Bougé, Jean-Marc Geib (équipe Espace, LIFL, Lille) et Brigitte Plateau (projet Apache, INRIA Rhône-Alpes) ont présenté le projet *iHPerf'98*, une *initiative informatique pour les hautes performances*. Il s'agissait de stimuler et de coordonner les efforts de *validation applicative* des recherches menées en France ces 10 dernières années dans ce domaine, en liaison étroite avec des partenaires industriels ou académiques non informaticiens. Une douzaine de *micro-actions* sont rassemblées. En novembre 1998, le MENRT a doté ce projet de 350 kF. À ce jour, 5 actions ont été ainsi soutenues.

GDR ARP, thème iHPerf. Le projet ReMaP participe activement aux activités du groupe thème iHPerf du GDR ARP (*Architectures, réseaux et systèmes, parallélisme*) sur l'algorithmique et les outils pour le parallélisme dans les applications régulières et irrégulières: L. Bougé, F. Desprez, J.-F. Méhaut, R. Namyst, L. Prylli, Y. Robert. URL: <http://www.prism.uvsq.fr/public/jfcollar/ihperf.html>.

GDR ARP, action Grappes. L. Bougé, J.-F. Méhaut, R. Namyst, L. Prylli et leurs thésards participent aux activités de cette action dirigée par Jean-Louis Pazat, Projet PARIS, IRISA. URL: <http://www.irisa.fr/grappes>.

GDR ARP, action TAROT. I. Guérin Lassous et N. Schabanel participent au groupe TAROT du GDR ARP sur les télécommunications. Ils y ont fait des exposés en octobre 1999. ReMaP apporte une compétence algorithmique dans ce groupe de recherche animé par Pierre Fraignaud du LRI. URL: <http://www.lri.fr/~pierre/TAROT>.

GDR ARP, thème Ordonnancement. Alain Darte participe aux activités du thème *Ordonnancement* du GDR ARP qui s'intéresse aux problèmes d'ordonnancement de toute nature et qui est dirigé par Philippe Chrétienne (LIP6).

ARC OURAGAN. Frédéric Desprez, Stéphane Domas, Martin Quinson, Cyril Randriamaro, Frédéric Suter font partie de l'action de recherche coopérative *OURAGAN* sur les outils pour la parallélisation de grands systèmes numériques. Cette action, créée pour 2 ans à l'hiver 1998, réunit les projets Métalau (Rocquencourt), ReMaP, Résédas (INRIA Lorraine), l'équipe ALIENOR du LaBRI (Université Bordeaux I), l'équipe Paladin (LaRIA, Université de Picardie, Amiens) et une équipe du LIFC (Université de Franche-Comté).

Les travaux effectués dans ReMaP concernent la parallélisation du logiciel Scilab et plus particulièrement l'automatisation de l'utilisation de bibliothèques numériques sur plates-formes de métacomputing et le développements de serveurs de calculs et outils logiciels associés.

ARC CoMPaS. I. Guérin Lassous fait partie de l'action de recherche coopérative (ARC) *CoM-PaS* sur les communications multi-points et qualité de service dans les réseaux locaux sans fil. Cette action est dirigée par Laurent Viennot, INRIA/Hipercom. Elle réunit les projets Hipercom (INRIA Rocquencourt), Résédas (INRIA Lorraine) et l'équipe Algorithmique et Combinatoire de Michel Morvan au LIAFA. URL: <http://menetou.inria.fr/compas>.

ACI Jeunes Chercheurs 2000. I. Guérin Lassous fait partie de l'Action concertée incitative *Jeunes chercheurs 2000* intitulée *Modélisation et optimisation de réseaux locaux sans fil* et dirigée par Stéphane Ubéda (CITI/INSA de Lyon). Cette action, créée pour 3 ans en septembre 2000 réunit le laboratoire CITI de l'INSA de Lyon, le projet Hipercom de l'INRIA Rocquencourt et le projet ReMaP. Cette action a pour but d'étudier et de proposer des outils d'aide à la planification de réseaux locaux sans fil.

8.2 Actions financées par la Commission européenne

Projet ProxiTV. Le projet ProxiTV (*A proxy infrastructure for Internet interactive TV*) est un projet financé par la communauté européenne qui répond à l'action IV du programme IST (Information Society Technology). Ce projet a commencé au 1^{er} octobre 2000 et dure deux ans. Les partenaires de ce projet sont Sycomore Aérospatiale Matra, Web-freeTV, Jet2web (filiale de Telekom Austria), Europsort et l'INRIA. Au sein de l'INRIA, Eric Fleury (INRIA Lorraine, projet Résédas) et Isabelle Guérin Lassous (INRIA Rhône-Alpes, projet ReMaP) participent à ce projet. ProxiTV a pour objectif de mettre en place une infrastructure proxy pour la télévision interactive sur Internet. Le rôle de l'INRIA est de proposer des algorithmes de routage dans un environnement hautement hétérogène de manière optimale. L'optimalité est définie comme une fonction de coût qui peut être basée sur le prix de la route choisie ou sur la longueur de la route choisie ou sur d'autres critères à définir par les partenaires. Un démonstrateur doit aussi être réalisé.

8.3 Réseaux et groupes de travail internationaux

Calcul creux et arithmétique d'intervalle. Nous avons participé à la mise en place d'une nouvelle collaboration avec Pierre Manneback, Faculté Polytechnique de Mons, Belgique, et Nathalie Revol, Laboratoire d'Analyse Numérique et d'Optimisation de Lille, et Mohamed Daoudi, Université d'Oujda au Maroc, autour de l'utilisation de PM2 pour des applications scientifiques à haute performance. Le projet DAPPI (Développement d'Algorithmes Parallèles pour des Problèmes Irréguliers) est soutenu par la commission européenne dans le cadre de son programme INCO-DC. Les problèmes irréguliers

étudiés dans ce projet concernent les domaines de l'algèbre linéaire creuse et l'optimisation globale par intervalles. F. Desprez et J.-F. Méhaut ont participé en octobre 1999 à l'animation de l'école ParDi '99 (<http://spirou.fpms.ac.be/pardi/index.html>) pour les jeunes chercheurs de l'Afrique du Nord intéressés par ce domaine de recherche. L'environnement PM2 a également été installé sur la grappe de PC du Laboratoire de Recherche en Informatique (LaRI) de l'Université d'Oujda.

Laboratoires de Hewlett Packard, USA. Alain Darte collabore avec Rob Schreiber et Bob Rau des laboratoires de HP à Palo Alto sur des problèmes de synthèse automatique de circuits. Cette collaboration a donné lieu à deux brevets sur des techniques de transformation et de génération de code. Guillaume Huard y a également travaillé 3 mois pendant l'été 2000 sur des optimisations et transformations de code.

8.4 Relations bilatérales internationales

Contrat NSF/INRIA, USA. L. Bougé, R. Namyst et J.-F. Méhaut collaborent avec l'équipe de Phil Hatcher et Robert Russell, Dept. Comp. Science, Univ. New Hampshire, Durham, NH, USA, dans le domaine des supports d'exécution pour les compilateurs parallèles. Phil Hatcher a en particulier développé un compilateur data-parallèle C* et un environnement d'exécution *distribué* pour Java. Cette collaboration a été soutenue par un contrat NSF/INRIA conclu à l'automne 1997 pour 2 ans. Il a obtenu une prolongation sans frais (*no-cost*) pour le premier semestre de l'année 2000. Un autre contrat est actuellement en cours de négociation sur la suite de ce travail. Il s'agit de mettre en œuvre l'environnement distribué *Hyperion* pour Java développé par Phil Hatcher au-dessus de la couche DSM (*Distributed Shared Memory*) de PM2. La réponse est attendue d'ici la fin de l'année.

Contrat UIUC/SPI CNRS, USA. L. Bougé, R. Namyst et J.-F. Méhaut collaborent avec l'équipe de Catherine Roucairol et Van-Dat Cung, PRiSM, Univ. Versailles Saint-Quentin et l'équipe de Laxmikant V. Kalé, Dept. Comp. Science, Univ. of Illinois at Urbana Champaign (UIUC). Cette collaboration concerne la mise en œuvre d'algorithmes haute performance d'optimisation combinatoire au-dessus de PM2 et du système Converse de Kalé. Elle est soutenue par un contrat de collaboration bilatérale conclu entre le département SPI du CNRS et l'UIUC à l'été 1998 pour 2 ans. La bibliothèque d'optimisation (BOB) développée par l'équipe de Versailles est aujourd'hui disponible sur les environnements PM2 et Converse. Des études comparatives d'évaluation de performance vont être menées cette année. La mise en place de liaisons transatlantiques performantes vont aussi nous permettre de mettre le développement d'applications d'optimisation combinatoire à plus large échelle (métacomputing). Ce contrat vient de se terminer.

Contrat LIAPUNOV INRIA/Université de Moscou, Russie. O. Beaumont, A. Legrand et Y. Robert collaborent avec l'Académie russe des Sciences dans le cadre d'un projet de l'Institut franco-russe A. M. Liapunov d'informatique et de mathématiques appliquées

sur la distribution des données sur des réseaux hétérogènes de stations pour des problèmes d'algèbre linéaire.

9 Diffusion de résultats

9.1 Animation de la communauté scientifique

Responsabilité d'animation

Département STIC du CNRS. Yves Robert a été nommé membre du conseil de département STIC du CNRS.

GDR CNRS ARP. L. Bougé a été nommé au 1^{er} janvier 2000 responsable du Groupement de recherche (GDR) CNRS *Architecture, réseaux et systèmes, parallélisme* (ARP). Il regroupe une petite centaine d'équipes de recherche française du monde académique. URL: <http://www.ens-lyon.fr/LIP/ARP>.

ASP MENRT iHPerf'98. L. Bougé est responsable avec Jean-Marc Geib et Brigitte Plateau de l'Action de soutien sur programme (ASP) intitulée *Initiative informatique 1998 pour les hautes performances* (iHPerf'98). Cette action a pour but la validation applicative des recherches faites ces dernières années dans le monde académique en informatique sur ce sujet, en étroite collaboration avec des utilisateurs finaux industriels ou académiques non informaticiens. Cette action est financée à hauteur de 350 kF par le MENRT à compter de novembre 1998.

RNTL. Luc Bougé a servi comme expert pour l'appel à propositions 2000 du RNTL.

Comités de rédaction, de pilotage et de programme

Yves Robert fait partie de l'Editorial Board de *Integration, the VLSI Journal* (North Holland) dans la section *Algorithms and Architectures*. Il est membre de l'Editorial Board de *Int. Journal Supercomputer Applications* (MIT Press). Il est l'éditeur européen de *Parallel Processing Letters* (World Scientific Publishing). Y. Robert a récemment été membre des comités de programme suivants : STACS 2000 (17th International Symposium on Theoretical Aspects of Computer Science), Lille ; ICPP 2000 (International Conference on Parallel Processing), Toronto (2000) – vice-président du comité de programme ; PDCS 2000 (13th International Conference on Parallel and Distributed Computing Systems), Las Vegas (2000) ; EuroPVM/MPI 2000, Lake Balaton, Hungary (2000).

Yves Robert a co-organisé l'an passé avec C. Mongenet, F. Vivien (ICPS Strasbourg) et S. Rajopadhye (IRISA Rennes) le colloque *Compilation et parallélisation automatique*. Ce colloque a été l'occasion de lancer un numéro spécial de la revue TSI (éditions Hermès) sur le thème *Techniques de parallélisation automatique*. Les articles soumis à ce numéro spécial sont actuellement en phase de révision.

Workshop MSA 2000. Le workshop *MSA 2000* (International Workshop on Metacomputing Systems and Applications) s'est tenu à Toronto en août 2000. Il a été co-organisé par Frédéric Desprez, Eric Fleury, Jean-François Méhaut et Yves Robert. Les actes sont publiés en conjonction avec ceux de ICPP 2000.

TSI. Alain Darté est membre du comité de rédaction de la revue *TSI* (Technique et science informatiques).

Workshop CPC 2000. Alain Darté, Yves Robert et Georges-André Silber ont organisé le dernier colloque CPC (Compilers for Parallel Computing) qui a eu lieu tous les 18 mois depuis 1989 : Oxford (1989), Paris (1990), Vienna (1992), Delft (1993), Malaga (1995), Aachen (1996), Vadstena (1998). Cette huitième édition a eu lieu à Aussois du 4 au 7 Janvier 2000 et a regroupé une cinquantaine de chercheurs essentiellement européens et américains de cette discipline, universitaires et industriels. Un numéro spécial de PPL (Parallel Processing Letters) a été organisé. Pour plus d'informations, voir <http://www.ens-lyon.fr/~darte/cpc2000/>.

Luc Bougé est vice-président du *Steering Committee* de la conférence internationale *Euro-Par* sur le parallélisme depuis 1995. Euro-Par 2000 avait lieu cette année à Munich, avec plus de 400 participants venus du monde entier. Il est coordinateur du *Comité de pilotage* des Rencontres francophones annuelles du parallélisme (*RenPar*). Depuis 1999, RenPar s'est associé à 2 autres conférences francophones : SympA (architecture) et CFSE (systèmes d'exploitation) pour organiser un événement commun annuel. Pour 2000, l'édition avait lieu à l'Université de Franche-Comté, Besançon. L'an prochain, il sera couplé aux Journées ASTI à La Villette (avril 2001). Luc Bougé a participé aux comités de programme suivants : *HiPC 2000*, 7th International Conference on High Performance Computing ; *PACT 2000*, International Conference on Parallel Architectures and Compilation Techniques ; *CLUSTER 2000*, IEEE International Conference on Cluster Computing ; *CMPP 2000*, 2nd International Workshop on Constructive Methods for Parallel Programming ; *RTSPP 2000*, 4th Workshop on Runtime Systems for Parallel Programming ; *PC-NOW 2000*, International Workshop on Personal Computer based Networks Of Workstations.

9.2 Enseignement universitaire

Responsabilités d'organisation

Magistère d'Informatique et Modélisation. R. Namyst est responsable de la 2^e année du magistère d'informatique et modélisation.

Concours d'entrée à l'ENS Lyon. Alain Darté est responsable de la filière informatique du concours d'entrée de l'École normale supérieure de Lyon.

École doctorale MathIF et DEA DIF. L. Bougé est responsable adjoint de l'École doctorale de mathématiques et d'informatique fondamentale (ED MathIF) créée à Lyon à la rentrée 1999. Cette école rassemble le DEA de mathématiques pures, le DEA d'analyse numérique et

le DEA d'informatique fondamentale (DEA DIF) qui fait suite au DIL en ce qui concerne le LIP et l'équipe ReSAM de B. Tourancheau.

Programme européen de l'ENS Lyon. L. Bougé a été co-responsable de ce programme pour l'informatique jusqu'à la rentrée 2000. Il s'agit de permettre les élèves 2^e année du Magistère d'informatique et de modélisation (MIM), donc du niveau maîtrise, de passer 6 mois dans une université européenne et de recevoir au MIM un nombre équivalent d'élèves européens.

Enseignement

DEA d'informatique de Lyon. En 2000-2001, plusieurs membres du projet enseignent au DEA d'informatique fondamentale (DIF): O. Beaumont (*Algorithmique parallèle et hétérogène*) et R. Namyst (*Supports d'exécution parallèle et distribuée*). URL: <http://www.ens-lyon.fr/DIF/>.

En 1999-2000, plusieurs membres du projet ont enseigné au nouveau DEA d'informatique fondamentale (DIF): A. Darte (*Ordonnancement et parallélisation automatique*), F. Desprez (*Algorithmique parallèle et parallélisation d'applications*), J.-F. Méhaut et R. Namyst (*Supports d'exécution parallèle et distribuée*), L. Prylli (*Communications dans les réseaux locaux et système d'exploitation*), S. Ubéda (*Optimisation des réseaux sans fil*). URL: <http://www.ens-lyon.fr/DIF/>.

DEA ID (Orsay). Raymond Namyst effectue la moitié d'un cours de DEA intitulé *Supports d'exécution parallèles et distribués* dans le DEA ID d'Orsay. L'autre moitié est assurée par Jean-François Méhaut, membre extérieur de ReMaP.

ENSEIRB, Bordeaux. Frédéric Desprez a donné un cours en troisième année à l'École Nationale Supérieure d'Électronique et de Radio-Électricité de Bordeaux (ENSERB) dans la filière informatique sur le placement de données et l'équilibrage de charge pour les applications numériques de grande taille.

DESS Réseau de Lyon 1. Frédéric Desprez a donné des cours et travaux pratiques sur MPI dans ce DESS.

9.3 Autres enseignements

Yves Robert est Maître de Conférences d'exercice partiel à l'École Polytechnique pour l'année 2000–2001.

Raymond Namyst a été invité à effectuer un cours intitulé *Introduction à la programmation à l'aide de processus légers* dans le cadre de l'école Jeunes Chercheurs organisée à l'ENS Lyon par le GDR Programmation.

Raymond Namyst a été invité à effectuer un tutoriel *Multithreading et calcul haute performance* à la conférence RenPar 2000 (conjointement avec Jean-François Méhaut).

Raymond Namyst a été invité à effectuer un cours sur les supports d'exécution parallèles à l'école iHPerf2000 (*applications parallèles haute performance: analyse, conception et utilisation*

de grappes homogènes ou hétérogènes de calculateurs) sur le thème des supports exécutifs pour grappes de PC.

Frédéric Desprez a été invité à donner un cours à SupElec dans le cadre de la formation continue sur les environnements pour la parallélisation d'applications numériques.

Frédéric Desprez a été invité à donner un cours sur la parallélisation d'applications numériques par passage de messages lors de l'École d'automne ParDi organisée à Oujda (Maroc) en octobre 1999.

Frédéric Desprez a été invité à donner un tutoriel sur MPI et OpenMP lors de la conférence Euro-Par '99 organisée à Toulouse en septembre 1999.

Olivier Beaumont, Frédéric Suter, Vincent Danjean et Olivier Aumage ont participé à l'accueil des classes à l'ENS-Lyon lors de la Semaine de la Fête de la Science.

Alain Darte a été examinateur de l'oral de mathématiques et informatique des concours d'entrée 1999 et 2000 de l'ENS-Lyon. Un recueil des exercices posés lors de ces deux sessions est en cours de publication.

Isabelle Guérin Lassous a été invitée à effectuer un cours intitulé *GPRS (General Packet Radio Services)* en cinquième année de la Formation d'ingénieur en informatique de la Faculté d'Orsay (FIIFO).

Luc Bougé a donné un cours sur la sémantique du parallélisme dans le cadre de l'école Jeunes Chercheurs organisée à l'ENS Lyon par le GDR Programmation. Il a donné un séminaire sur le *metacomputing* aux élèves du Magistère d'informatique de l'ENS Lyon.

Loïc Prylli a été invité à donner un séminaire au DEA d'Informatique de Grenoble.

9.4 Participation à des colloques, séminaires, invitations

Yves Robert a été invité à donner des séminaires au Computer Science Department de l'Université de Pittsburgh, et au séminaire Algorithmique du LIX/LRI. Il a été invité à donner une conférence tutorielle *Algorithmique pour réseaux hétérogènes* à la conférence RenPar 2000.

Yves Robert a participé à la table ronde *La gestion de la R&D dans les entreprises : fonction interne ou externe?* organisée par les entretiens du Grand Lyon Technopole.

Frédéric Desprez a été invité à présenter le projet OURAGAN dans le cadre de la conférence CCGSC 2000 qui s'est déroulée à Faverges en Septembre 2000.

Nicolas Schabanel a donné une conférence invitée *A state of the Art in Data Dissemination* au séminaire de Dagstuhl *Linear, Semidefinite Programming and Randomization Methods for Combinatorial Optimization Problems* organisé par K. Jansen (Kiel), J. Rolim (Geneva) et M. Sudan (MIT Cambridge), en janvier 2000.

Raymond Namyst a accordé une interview à la chaîne télévisée Lyonnaise *Cité Télévision* à propos du transfert technologique au LIP, et plus particulièrement à propos du projet MOB3D qu'il a supervisé l'année dernière.

Alain Darte a participé au colloque *Compilation et Parallélisation Automatique* organisé par l'ICPS en octobre 1999.

Alain Darte a donné un tutoriel et un exposé plus pointu en novembre 1999 à Tunis dans le cadre d'une semaine sur le parallélisme.

Alain Darte a fait un séminaire au LIRMM sur le décalage d'instructions pour la compaction de boucles sur invitation de Jean-Claude Bajard et Jean-Claude König.

Georges-André Silber a fait un séminaire à l'Université de Marseille sur la parallélisation automatique par insertions de directives en novembre 1999.

Guillaume Huard a fait un séminaire sur le décalage d'instructions pour la compaction de boucles à l'INRIA Rocquencourt en octobre 1999, un exposé sur le décalage d'instructions pour la détection de boucles parallèles à l'ENSMP (dans le cadre de l'action *iHPerf*) en mars 2000.

Isabelle Guérin Lassous a donné des séminaires sur les *Modèles parallèles à gros grain et leurs applications sur les graphes* à l'INRIA Rocquencourt, au LARIA à Amiens et au laboratoire ID à Grenoble. Elle a été invitée à présenter les *Réseaux ad hoc* lors d'une journée de recherche en télécommunications au laboratoire CITI de l'INSA de Lyon.

Olivier Aumage a fait un exposé au séminaire du laboratoire d'informatique de l'Université du New Hampshire (USA), dans le cadre d'une réunion *Grappes* (action transversale du GdR ARP), dans le cadre de la semaine sport-étude du MIM2, lors du groupe de travail Calcul parallèle et distribué du LaBRI.

Luc Bougé a été invité à participer au séminaire de Dagstuhl *High Performance Computing and Java* en août 2000. Il a donné un exposé sur le projet Hyperion d'environnement Java distribué mené en collaboration avec Phil Hatcher, UNH. Il a été invité à présenter les travaux récents effectués sur PM2 dans le cadre de la conférence CCGSC 2000 qui s'est déroulée à Faverges en septembre 2000.

Luc Bougé a présenté les travaux réalisés autour des serveurs multimédia dans le LHPC lors de la présentation des travaux de l'INRIA aux responsables de Thomson Multimédia à Rocquencourt, en septembre 2000.

Luc Bougé a été interviewé par *le monde informatique* lors de la conférence Euro-Par '99 à Toulouse (parution dans le numéro 828 du 5 novembre 1999).

Loïc Prylli a été invité à donner une présentation sur les grappes de PC Linux au journées Linux organisé à Mons, Belgique.

Loïc Prylli a participé au reportage de TLM (Télé Lyon Métropole) pour l'émission Télé-Campus.

10 Bibliographie

Livres et monographies

- [1] A. DARTE, Y. ROBERT, F. VIVIEN, *Scheduling and Automatic Parallelization*, Birkhauser, 2000, ISBN 0-8176-4149-1.
- [2] A. DARTE, Y. ROBERT (éditeurs), *Special issue on Compilers for parallel computers*, Parallel Processing Letters, 2000. À paraître.
- [3] C. MONGENET, Y. ROBERT (éditeurs), *Special issue on Techniques de parallélisation automatique*, Techniques et Science Informatiques, 2000. À paraître.

Thèses et habilitations à diriger des recherches

- [4] A. DARTE, *De l'organisation des calculs dans les codes répétitifs*, thèse de doctorat, Université Claude Bernard de Lyon, décembre 1999, Disponible comme rapport LIP HDR1999-03.

- [5] J.-F. MÉHAUT, *Supports d'exécution pour le calcul parallèle et distribué*, Habilitation à diriger les recherches, Univ. Claude Bernard, Lyon, LIP, ENS Lyon, janvier 2000, Disponible comme rapport LIP HDR2000-01, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/HDR/HDR2000/HDR2000-01.ps.Z>.
- [6] C. PEREZ, *Compilation des langages à parallélisme de données : gestion de l'équilibrage de charge par un exécutif à base de processus légers*, Thèse de doctorat, ENS Lyon, France, LIP, décembre 1999, Disponible comme rapport LIP PhD1999-10, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/PhD/PhD1999/PhD1999-10.ps.Z>.
- [7] F. RASTELLO, *Partitionnement : optimisations de compilation et algorithmique hétérogène*, thèse de doctorat, École Normale Supérieure de Lyon, septembre 2000.
- [8] N. SCHABANEL, *Algorithmes d'approximation pour les télécommunications sans fil : Dissémination de données et Allocation statique de fréquences*, thèse de doctorat, ENS-Lyon, janvier 2000.
- [9] G.-A. SILBER, *Parallélisation automatique par insertion de directives*, thèse de doctorat, ENS Lyon, décembre 1999, Disponible comme rapport LIP PhD1999-12.

Articles et chapitres de livre

- [10] G. ANTONIU, L. BOUGÉ, R. NAMYST, C. PEREZ, «Compiling Data-parallel Programs to A Distributed Runtime Environment with Thread Isomigration», *Parallel Processing Letters*, mars 2000, À paraître. Special issue on CPC 2000., <http://www.ens-lyon.fr/~bouge/Biblio/Perez/AntBouNamPer00PPL.ps.gz>.
- [11] P. BOULET, J. DONGARRA, F. RASTELLO, Y. ROBERT, F. VIVIEN, «Algorithmic issues on heterogeneous computing platforms», *Parallel Processing Letters* 9, 2, 1999, p. 197–213.
- [12] F. CHAUSSUMIER, F. DESPREZ, «Communications Optimizations and Efficient Load-Balancing for a Volume Rendering Algorithm on a Cluster of PCs», *Parallel and Distributed Computing Practices*, 2000, À paraître.
- [13] A. DARTE, G. HUARD, «Loop Shifting for Loop Compaction», *International Journal of Parallel Programming* 28, 5, 2000, p. 499–534.
- [14] A. DARTE, Y. ROBERT, F. VIVIEN, «Loop Parallelization Algorithms», in : *Languages, Compilers and Run Time Systems for Scalable Parallel Systems*, S. Pande (éditeur), *Lecture Notes in Computer Science*, 1808, Springer Verlag, 2000.
- [15] A. DARTE, «On the Complexity of Loop Fusion», *Parallel Computing* 26, 9, juillet 2000, p. 1175–1193.
- [16] A. FERREIRA, N. SCHABANEL, «A randomized BSP algorithm for the maximal independent set problem.», *Parallel Processing Letters*, 2000, À paraître.

Communications à des congrès, colloques, etc.

- [17] G. ANTONIU, L. BOUGÉ, P. HATCHER, M. MACBETH, K. MCGUIGAN, R. NAMYST, «Compiling multithreaded Java bytecode for distributed execution», in : *Euro-Par 2000: Parallel Processing, Lect. Notes in Comp. Science*, 1900, Springer-Verlag, p. 1039–1052, Munchen, Germany, août 2000. Selected as a Distinguished Paper, <http://www.ens-lyon.fr/~bouge/Biblio/Antoniou/AntBouHatBetGuiNam00EuroPar.ps.gz>.

- [18] G. ANTONIU, L. BOUGÉ, P. HATCHER, M. MACBETH, K. MCGUIGAN, R. NAMYST, «Implementing Java consistency using a generic, multithreaded DSM runtime system», in : *Parallel and Distributed Processing. Proc. Intl Workshop on Java for Parallel and Distributed Computing, Lect. Notes in Comp. Science, 1800*, Held in conjunction with IPDPS 2000. IEEE TCPP and ACM, Springer-Verlag, p. 560–567, Cancun, Mexico, mai 2000, <http://www.ens-lyon.fr/~bougé/Biblio/Antoniou/AntBouHatBetGuiNam00JavaPDC.ps.gz>.
- [19] G. ANTONIU, L. BOUGÉ, R. NAMYST, «DSM-PM2: a multi-protocol DSM layer for the PM2 multithreaded runtime system», in : *Proc. 2nd Workshop on Parallel Computing for Irregular Applications (WPCIA2)*, J. Chassin de Kergommeaux, B. Lecussan, J.-F. Méhaut (éditeurs), Toulouse, France, janvier 2000. Held in conjunction with HPCA-6. Internal proceedings only, <http://www.ens-lyon.fr/~bougé/Biblio/Antoniou/AntBouNam00WPCIA2.ps.gz>.
- [20] O. AUMAGE, L. BOUGÉ, A. DENIS, J.-F. MÉHAUT, G. MERCIER, R. NAMYST, L. PRYLLI, «A Portable and Efficient Communication Library for High-Performance Cluster Computing», in : *IEEE Intl Conf. on Cluster Computing (CLUSTER 2000)*, p.??, Technische Universität Chemnitz, Saxony, Germany, novembre 2000. À paraître, <http://www.ens-lyon.fr/~bougé/Biblio/Aumage/AumBouDenMehMerNamPry00Cluster2000.ps.gz>.
- [21] O. AUMAGE, L. BOUGÉ, R. NAMYST, «A Portable and Adaptative Multi-Protocol Communication Library for Multithreaded Runtime Systems», in : *Parallel and Distributed Processing. Proc. 4th Workshop on Runtime Systems for Parallel Programming (RTSPP '00)*, Lect. Notes in Comp. Science, 1800, Held in conjunction with IPDPS 2000. IEEE TCPP and ACM, Springer-Verlag, p. 1136–1143, Cancun, Mexico, mai 2000, <http://www.ens-lyon.fr/~bougé/Biblio/Aumage/AumBouNam00RTSPP.ps.gz>.
- [22] O. AUMAGE, «Communications multiprotocoles sur grappes de stations : l'approche Madeleine II», in : *Actes des Rencontres francophones du parallélisme (RenPar 12)*, p. 73–78, LIB, Univ. Besançon, juin 2000. Présentation sélectionnée pour publication dans le numéro spécial RenPar 12 de TSI, <http://www.ens-lyon.fr/~bougé/Biblio/Aumage/Aum00RenPar.ps.gz>.
- [23] O. BEAUMONT, V. BOUDET, A. LEGRAND, F. RASTELLO, Y. ROBERT, «Heterogeneity Considered Harmful to Algorithm Designers», in : *Cluster'2000*, IEEE Computer Society Press, 2000.
- [24] O. BEAUMONT, V. BOUDET, A. LEGRAND, F. RASTELLO, Y. ROBERT, «Heterogeneous Matrix-Matrix Multiplication, or Partitioning a Square into Rectangles : NP-Completeness and Approximation Algorithms», in : *EuroMicro Workshop on Parallel and Distributed Computing (EuroMicro'2001)*, IEEE Computer Society Press, 2001.
- [25] O. BEAUMONT, V. BOUDET, F. RASTELLO, Y. ROBERT, «Load balancing strategies for dense linear algebra kernels on heterogeneous two-dimensional grids», in : *14th International Parallel and Distributed Processing Symposium (IPDPS'2000)*, IEEE Computer Society Press, p. 783–792, 2000.
- [26] O. BEAUMONT, V. BOUDET, F. RASTELLO, Y. ROBERT, «Matrix-matrix multiplication on heterogeneous platforms», in : *2000 International Conference on Parallel Processing (ICPP'2000)*, IEEE Computer Society Press, p. 289–298, 2000.
- [27] K. BERTET, I. GUÉRIN LASSOUS, L. VIENNOT, «Un premier pas vers la réservation de bande passante dans les réseaux radio», in : *Actes des 2es Rencontres francophones sur les aspects algorithmiques des télécommunications (Algotel 2000)*, 2000.

- [28] A. BONHOMME, P. GEOFFRAY, «High Performance Video Server using Myrinet», in : *Myrinet User Group Conference*, p. 66–72, Lyon, France, septembre 2000, <http://www.ens-lyon.fr/~bougé/Biblio/Bonhomme/BonGeo00MUG.ps.gz>.
- [29] A. BONHOMME, L. PRYLLI, «A distributed storage system for a video-on-demand server», in : *Euro-Par 2000: Parallel Processing, Lect. Notes in Comp. Science, 1900*, Springer-Verlag, p. 1110–1114, Munchen, Germany, août 2000, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/RR/RR2000/RR2000-16.ps.Z>.
- [30] A. BONHOMME, L. PRYLLI, «Un système de stockage distribué pour un serveur de vidéo à la demande (version révisée)», in : *3e colloque international sur les nouvelles technologies de la répartition (NOTERE 2000)*, Paris, France, novembre 2000. Version révisée et étendue de [29]. À paraître., <http://www.ens-lyon.fr/~bougé/Biblio/Bonhomme/BonPry00NOTERE.ps.gz>.
- [31] V. BOUDET, F. RASTELLO, Y. ROBERT, «PVM implementation of heterogeneous ScaLAPACK dense linear solvers», in : *Recent Advances in Parallel Virtual Machine and Message Passing Interface*, J. Dongarra, E. Luque, T. Margalef (éditeurs), LNCS, 1697, Springer Verlag, p. 333–340, 1999.
- [32] L. BOUGÉ, J.-F. MÉHAUT, R. NAMYST, L. PRYLLI, «Using the VI Architecture to build distributed, multithreaded runtime systems: a case study», in : *Proc. 2000 ACM Symposium on Applied Computing (SAC 2000)*, ACM Special Interest Group on Applied Computing (SIGAPP), ACM, p. 704–709, Villa Olmo, Como, Italy, mars 2000, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/RR/RR1999/RR1999-27.ps.Z>.
- [33] F. CHAUSSUMIER, F. DESPREZ, M. LOI, «Efficient Load-Balancing and Communication Overlap in Parallel Shear-Warp Algorithm on a Cluster of PCs», in : *Proceedings of EuroPAR'99*, P. Amestoy, P. Berger, M. Daydé, I. Duff, V. Frayssé, L. Giraud, D. Ruiz (éditeurs), *Lecture Notes in Computer Science*, 1685, Springer Verlag, p. 570–577, Toulouse, 1999.
- [34] F. CHAUSSUMIER, F. DESPREZ, L. PRYLLI, «Asynchronous Communications in MPI – the BIP/Myrinet Approach», in : *Proceedings of the EuroPVM/MPI'99 conference*, J. Dongarra, E. Luque, T. Margalef. (éditeurs), *Lecture Notes in Computer Science*, 1697, Springer Verlag, p. 485–492, Barcelona, Spain, September 1999.
- [35] V. DANJEAN, R. NAMYST, R. RUSSELL, «Integrating Kernel Activations in a Multithreaded Runtime System on Linux», in : *Parallel and Distributed Processing. Proc. 4th Workshop on Runtime Systems for Parallel Programming (RTSPP '00), Lect. Notes in Comp. Science, 1800*, Held in conjunction with IPDPS 2000. IEEE TCPP and ACM, Springer-Verlag, p. 1160–1167, Cancun, Mexico, mai 2000, <http://www.ens-lyon.fr/~bougé/Biblio/Danjean/DanNamRus00RTSPP.ps.gz>.
- [36] V. DANJEAN, R. NAMYST, R. RUSSELL, «Linux Kernel Activations to Support Multithreading», in : *Proc. 18th IASTED International Conference on Applied Informatics (AI 2000)*, IASTED, p. 718–723, Innsbruck, Austria, février 2000, <http://www.ens-lyon.fr/~bougé/Biblio/Danjean/DanNamRus00IASTED.ps.gz>.
- [37] V. DANJEAN, «LinuxActivations: un support système performant pour les applications de calcul multithreads», in : *Actes des Rencontres francophones du parallélisme (RenPar 12)*, p. 87–92, LIB, Univ. Besançon, juin 2000, <http://www.ens-lyon.fr/~bougé/Biblio/Danjean/Dan00RenPar.ps.gz>.

-
- [38] A. DARTE, C. DIDERICH, M. GENGLER, F. VIVIEN, «Scheduling the Computations of a Loop Nest with Respect to a Given Mapping», in : *European Conference on Parallel Computing (Euro-Par'00)*, München, Germany, août 2000.
- [39] A. DARTE, R. SCHREIBER, B. R. RAU, F. VIVIEN, «A Constructive Solution to the Juggling Problem in Systolic Array Synthesis», in : *Proceedings of the International Parallel and Distributed Processing Symposium (IPDPS'00)*, p. 815–821, Cancun, Mexico, mai 2000.
- [40] A. DARTE, G.-A. SILBER, «Temporary Arrays for Distribution of Loops with Control Dependence», in : *European Conference on Parallel Computing (Euro-Par'00)*, München, Germany, août 2000.
- [41] A. DARTE, «On the Complexity of Loop Fusion», in : *International Conference on Parallel Architectures and Compilation Techniques (PACT'99)*, p. 149–157, octobre 1999.
- [42] A. DENIS, «Adaptation de l'environnement générique de métacomputing Globus à des réseaux hauts débits», in : *Actes des Journées Doctorales Informatique et Réseaux (JDIR 2000)*, Univ. Paris 6, novembre 2000, <http://www.ens-lyon.fr/~bougue/Biblio/Denis/Den00JDIR.ps.gz>.
- [43] A. DENIS, «VRP: un protocole avec une tolérance de perte ajustable pour des hautes performances sur réseau longue distance», in : *Actes des Rencontres francophones du parallélisme (Ren-Par 12)*, p. 27–32, LIB, Univ. Besançon, juin 2000, <http://www.ens-lyon.fr/~bougue/Biblio/Denis/Den00RenPar.ps.gz>.
- [44] F. DESPREZ, S. DOMAS, «Efficient Pipelining of Level 3 BLAS Routines», in : *4th international meeting VECPAR 2000*, 3, p. 675–688, Porto, June 2000.
- [45] F. DESPREZ, C. RANDRIAMARO, «Redistribution entre appels de routines ScaLAPACK», in : *Actes des 12èmes Rencontres Francophones du Parallélisme RenPar'00*, Besançon, France, June 2000.
- [46] A. GEBREMEDHIN, I. GUÉRIN LASSOUS, J. GUSTEDT, J. TELLE, «Graph Coloring for a Coarse Grained Multiprocessor», in : *Proceedings of Workshop on Graph-Theoretical Concepts (WG 2000)*, U. Brandes, D. Wagner (éditeurs), LNCS, 1928, Springer Verlag, p. 184–195, 2000.
- [47] P. GEOFFRAY, L. PRYLLI, B. TOURANCHEAU, «BIP-SMP: High Performance message passing over a cluster of commodity SMPs», in : *Supercomputing (SC '99)*, Portland, OR, novembre 1999. Electronic proceedings only, available at <http://www.supercomp.org/sc99/proceedings/>, <http://www.supercomp.org/sc99/proceedings/papers/geoffray.pdf>.
- [48] I. GUÉRIN LASSOUS, J. GUSTEDT, M. MORVAN, «Graphs According a Coarse Grained Approach: Experiments with PVM and MPI», in : *Proceedings of European PVM/MPI Users' Group Meeting (EuroPVM/MPI 2000)*, J. Dongarra, P. Kacsuk, N. Podhorszki (éditeurs), LNCS, 1908, Springer Verlag, p. 72–79, 2000.
- [49] I. GUÉRIN LASSOUS, J. GUSTEDT, «Portable List Ranking: an Experimental Study», in : *Proceedings of Workshop on Algorithm Engineering (WAE 2000)*, LNCS, Springer Verlag, 2000. À paraître.
- [50] I. GUÉRIN LASSOUS, E. THIERRY, «Generating Random Permutations in the Parallel Coarse Grained Models Framework», in : *Proceedings of the International Conference on Principles of Distributed Systems (OPODIS 2000)*, 2000. À paraître.

- [51] C. KENYON, N. SCHABANEL, N. YOUNG, «Polynomial-Time Approximation Scheme for Data Broadcast», in : *Proc. of the 32nd ACM Symp. on Theory of Computing (STOC 2000)*, p. 659–666, 2000.
- [52] A. PETITET, H. CASANOVA, J. DONGARRA, Y. ROBERT, R. WHALEY, «Parallel and Distributed Scientific Computing: A Numerical Linear Algebra Problem Solving Environment Designer's Perspective», in : *Handbook on Parallel and Distributed Processing*, J. Blazewicz, K. Ecker, B. Plateau, D. Trystram (éditeurs), *International Handbook on Information Systems*, 3, Springer Verlag, p. 464–504, 2000. Available as LAPACK Working Note 139.
- [53] L. PRYLLI, R. WESTRELIN, «Current issues in available implementations on Myrinet», in : *Proc. First Myrinet User Group Conference*, p. 149–155, Lyon, France, septembre 2000, <http://www.ens-lyon.fr/~bougé/Biblio/Prylli/PryWes00MUG.ps.gz>.
- [54] N. SCHABANEL, «The databroadcast problem with preemption», in : *Proc. of the 17th International Symposium on Theoretical Aspects of Computer Science (STACS 2000)*, LNCS, 1770, Springer Verlag, p. 181–192, 2000.
- [55] F. SUTER, E. CARON, D. LAZURE, «Manipulation de données de grande taille dans Scilab//», in : *Douzièmes Rencontres Francophones du Parallélisme*, Besançon, 19-22 juin 2000.

Rapports de recherche et publications internes

- [56] G. ANTONIU, L. BOUGÉ, R. NAMYST, «Generic Distributed Shared Memory: the DSM-PM2 Approach», *Research Report n° RR2000-19*, LIP, ENS Lyon, Lyon, France, mai 2000, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/RR/RR2000/RR2000-19.ps.Z>.
- [57] O. AUMAGE, G. MERCIER, R. NAMYST, «MPICH/Madeleine: A True Multi-Protocol MPI for High-Performance Networks», *Research Report n° RR2000-30*, LIP, ENS Lyon, Lyon, France, octobre 2000, Also available as Research Report RR-4016, INRIA Rhône-Alpes. Submitted for presentation at IPDPS 2001, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/RR/RR2000/RR2000-30.ps.Z>.
- [58] L. BOUGÉ, J.-F. MÉHAUT, R. NAMYST, «Madeleine: An Efficient and Portable Communication Interface for RPC-based Multithreaded Environments (revised version)», *Research report*, LIP, ENS Lyon, décembre 1999, Submitted for publication in *Journal of Parallel and Distributed Computing (JPDC)*, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/RR/RR1999/RR1999-55.ps.Z>.
- [59] F. CHAUSSUMIER, F. DESPREZ, M. LOI, «Efficient Load-balancing and Communication Overlap in Parallel Shear-Warp Algorithm on a Cluster of PCs», *rapport de recherche n° RR1999-28*, LIP ENS Lyon, 1999.
- [60] I. GUÉRIN LASSOUS, J. GUSTEDT, M. MORVAN, «Feasibility, Portability, Predictability and Efficiency: Four Ambitious Goals for the Design and Implementations of Parallel Coarse Grained Graph Algorithms», *rapport de recherche*.
- [61] G. MERCIER, «Ch_mad: un support efficace pour l'hétérogénéité des réseaux dans MPI», *Rapport de stage de DEA n° DEA2000-01*, LIP, ENS Lyon, Lyon, France, juin 2000, DEA d'informatique fondamentale, Univ. Claude Bernard, Lyon 1. Disponible en anglais comme [57], <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/DEA/DEA2000/DEA2000-01.ps.Z>.

- [62] J.-C. MIGNOT, «État de l'art des caches Web», *Research Report n° RR1999-52*, LIP, ENS Lyon, Lyon, France, décembre 1999, Rapport de contrat CHARM. Aussi disponible comme Rapport de recherche RR-3854 de l'INRIA Rhône-Alpes, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/RR/RR1999/RR1999-52.ps.Z>.
- [63] J.-C. MIGNOT, «Compression vidéo et nouveaux codecs : un état de l'art», *Rapport de contrat SPIHD*, LIP, ENS Lyon, novembre 2000, À paraître.
- [64] J.-C. MIGNOT, «Distribution de contenus multimédia par flot continu : état de l'art», *Rapport de contrat CHARM*, LIP, ENS Lyon, octobre 2000.
- [65] A. MOSTEFAOUI, A. BONHOMME, «Serveurs vidéo : concepts de base et prototypes», *Research Report n° RR1999-56*, LIP, ENS Lyon, Lyon, France, décembre 1999, Aussi disponible comme Rapport de recherche RR-3837 de l'INRIA Rhône-Alpes, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/RR/RR1999/RR1999-56.ps.Z>.
- [66] N. SCHABANEL, «The databroadcast problem with preemption», *rapport de recherche n° RR1999-49*, LIP, ENS Lyon, septembre 1999, <http://www.ens-lyon.fr/LIP/>.

Divers

- [67] G. ANTONIU, O. AUMAGE, L. BOUGÉ, L. LEFÈVRE, J.-F. MÉHAUT, R. NAMYST, C. PEREZ, «ARC INRIA "Réseaux à capacité d'adressage" (ResCapA) : rapport final de l'équipe lyonnaise», Document interne, février 2000, http://www.ens-lyon.fr/~bouge/Biblio/Bouge/ResCapA_final.ps.gz.
- [68] V. BERNARDI, *Introduction de protocoles de cohérence relâchée dans le système de mémoire virtuellement partagée DSM-PM2*, Rapport de stage MIM1, Magistère d'informatique et modélisation (MIM), ENS Lyon, LIP, ENS Lyon, septembre 2000, Stage effectué sous la direction de Gabriel Antoniu et Raymond Namyst, http://www.ens-lyon.fr/~bouge/Biblio/Bernardi/stage_MIM1.ps.gz.
- [69] L. BOUGÉ, «La programmation parallèle recherche la transparence», Interview dans *Le Monde Informatique*, numéro 828 du 5 novembre 1999, par Frédéric Noailles, novembre 1999, http://www.ens-lyon.fr/~bouge/Biblio/Bouge/article_LMI.html.
- [70] L. BOUGÉ, «Metacomputing : si tous les ordinateurs du monde...», Exposé au séminaire du Magistère d'informatique et modélisation (MIM), ENS Lyon, octobre 2000, <http://www.ens-lyon.fr/~bouge/Biblio/Bouge/Bou00Metacomputing.ppt>.
- [71] V. DANJEAN, *Environnement multithreads distribués : traitement efficace et réactif des communications*, Rapport de stage de DEA, DEA d'informatique fondamentale, Univ. Claude Bernard, Lyon 1, France, juin 2000, <ftp://ftp.ens-lyon.fr/pub/LIP/Rapports/DEA/DEA2000/DEA2000-03.ps.Z>.
- [72] A. DENIS, *Adaptation de l'environnement générique de metacomputing Globus à des réseaux haut débit*, Rapport de stage de DEA, DEA d'informatique fondamentale, Univ. Claude Bernard, Lyon 1, France, juin 2000, http://www.ens-lyon.fr/~bouge/Biblio/Denis/rapport_DEA.ps.gz.

- [73] L. EYRAUD, *Interconnexion efficace de réseaux haut débit: étude et réalisation d'une passerelle multiprotocole portable*, Rapport de stage MIM1, Magistère d'informatique et modélisation (MIM), ENS Lyon, LIP, ENS Lyon, septembre 2000, Stage effectué sous la direction de Olivier Aumage et Raymond Namyst, http://www.ens-lyon.fr/~bouge/Biblio/Eyraud/stage_MIM1.ps.gz.
- [74] R. NAMYST, J.-F. MÉHAUT, «Multithreading et calcul à haute performance», Tutoriel présenté à RenPar 2000, juin 2000, <http://www.ens-lyon.fr/~bouge/Biblio/Namyst/NamMeh00RenPar.ppt>.
- [75] R. NAMYST, «Le transfert technologique au LIP, ENS Lyon», Interview télévisée sur la chaîne câblée "Cités Télévision", le 10 octobre 2000, par Lionel Joly, octobre 2000.
- [76] N. VIOLLET, *Système de nommage et d'accès aux ressources logicielles dans un environnement de métacomputing*, Rapport de stage de DEA, DEA d'informatique fondamentale, Univ. Claude Bernard, Lyon 1, France, juin 2000, http://www.ens-lyon.fr/~bouge/Biblio/Viollet/rapport_DEA.ps.gz.