

Pierre Ramet

Curriculum Vitæ

Membre du projet HiePACS Inria Bordeaux Sud-Ouest et du LaBRI

Né le 13 octobre 1971 à Montreuil (93), marié.

Adresse personnelle : 1c rue du clos du Haut Carré, 33400 Talence.

Adresse professionnelle : Université Bordeaux, 351 cours de la Libération, 33405 Talence Cedex.

Formation

- 2016/2017 ◇ Habilitation à Diriger des Recherches, soutenue le 27 novembre 2017.
Sujet : *heterogeneous architectures, Hybrid methods, Hierarchical matrices for Sparse Linear Solvers*
- 1997/2000 ◇ Thèse de doctorat d'Informatique de l'Université Bordeaux sous la direction de F. Desprez et J. Roman, soutenue le 12 janvier 2000.
Sujet : *Optimisation de la communication et de la distribution des données pour des solveurs parallèles directs en algèbre linéaire dense et creuse.*
- 1996/1997 ◇ Service National effectué en tant que Scientifique du Contingent au CEA/CESTA.
Sujet : *Parallélisation et optimisation d'un code d'électromagnétisme.*
- 1994/1995 ◇ DEA d'Informatique de l'Université Bordeaux.
Sujet : *Calcul de la taille optimale de paquets pour les algorithmes macro-pipelines.*
- 1992/1995 ◇ Elève ingénieur à l'ENSEIRB, filière Informatique, option Parallélisme.

Activités d'enseignement et de recherche

- 2015/- ◇ Conseiller scientifique au CEA CESTA pour le domaine du HPC.
- 2000/- ◇ Maître de conférences au Département Informatique de l'IUT de l'Université Bordeaux. Mes activités de recherche sont actuellement rattachées au projet HiePACS Inria Bordeaux Sud-Ouest et du LaBRI (UMR 5800).
- 2011/2012 ◇ Délégation Inria dans le projet BACCHUS.
- 2003/2005 ◇ Délégation Inria dans le projet ScAlApplix.

Activités administratives

- 2011/- ◇ Responsable adjoint pour l'équipe SATANAS du LaBRI.
- 2011/2012 ◇ Responsable permanent pour le projet BACCHUS Inria.
- 2010/2017 ◇ Membre du bureau du MCIA (Mésocentre de Calcul Intensif Aquitain).
- 2009/2017 ◇ Membre de la Commission Consultative pour la 27ième section de l'Université Bordeaux.
- 2008/2016 ◇ Expert au GENCI pour le comité thématique 6 (Maths-Info).

- 2007/2014 ◇ Directeur des études des licences professionnelles SIL (Systèmes Informatiques et Logiciels) de l'Université Bordeaux.
- 2005/2008 ◇ Membre de la Commission de Spécialistes pour la 27ième section de l'Université Bordeaux et président du collège B en 2008.

Collaborations nationales

- 2013/2017 ◇ Membre du projet SOLHAR, ANR MN13.
- 2011/2015 ◇ Responsable scientifique Inria du projet ANEMOS, ANR MN11.
- 2008/2012 ◇ Membre des projets PETAL et PETALH, ANR COSINUS08 et COSINUS11.
- 2006/2010 ◇ Responsable scientifique Inria du projet ASTER, ANR CIS06.
- 2005/2009 ◇ Membre des projets NUMASIS et SOLTICE, ANR CIS05 et CIS06.

Collaborations internationales

- 2013/- ◇ Université de Stanford dans le cadre de l'équipe associée FAST-LA.
- 2011/- ◇ Université du Tennessee dans le cadre de l'équipe associée MORSE.
- 2008/2009 ◇ Université du Minnesota dans le cadre de l'équipe associée PHYLEAS.
- 2007/2008 ◇ Japan Atomic Energy Agency.

Encadrement doctoral en cours (6 thèses soutenues durant les 6 dernières années)

- 2015/2018 ◇ Co-encadrement de la thèse de Grégoire Pichon, contrat DGA-Inria.

Logiciels

PaStiX est un solveur haute performance pour la résolution de grands systèmes linéaires basé sur une approche directe supernodale. Ce solveur est reconnu par la communauté pour être un des plus performants sur les architectures récentes constituées de nœuds multicœurs. (> 120000 lignes de code, > 60000 téléchargements)

Publications (voir <http://www.labri.fr/perso/ramet/bib/Author/RAMET-P.html>)

5 publications marquantes :

1. G. Pichon, E. Darve, M. Faverge, P. Ramet, and J. Roman. *Sparse Supernodal Solver Using Block Low-Rank Compression : design, performance and analysis*. International Journal of Computational Science and Engineering, 27 :255-270, 2018.
2. G. Pichon, M. Faverge, P. Ramet, and J. Roman. *Reordering strategy for blocking optimization in sparse linear solvers*. SIAM Journal on Matrix Analysis and Application, 38(1) :226-248, 2017.
3. M. Barrault, B. Lathuilière, P. Ramet et J. Roman. *Efficient Parallel Resolution of The Simplified Transport Equations in Mixed-Dual Formulation*. Journal of Computational Physics, 230(5) :2004-2020, 2011.
4. P. Hénon, P. Ramet et J. Roman. *On finding approximate supernodes for an efficient ILU(k) factorization*. Parallel Computing, 34 :345-362, 2008.
5. P. Hénon, P. Ramet, et J. Roman. *PaStiX : A High-Performance Parallel Direct Solver for Sparse Symmetric Definite Systems*. Parallel Computing, 28(2) :301-321, 2002.

Activités d'enseignement

J'effectue mon service statutaire au Département Informatique de l'IUT de l'Université Bordeaux 1. Je dispense, en particulier, des cours en Programmation, Bases de Données, Systèmes et Réseaux. Depuis la rentrée 2005, je suis impliqué dans le cadre de la licence professionnelle (Systèmes Informatiques et Logiciels) en assurant les cours de Bases de Données Avancées de Programmation Orientée Objet et d'Objets Distribués.

Je dispense également des cours (Bases de Données Parallèles et Ordonnancement et Algèbre Linéaire) en 3^{ième} année à l'ENSEIRB dans la filière informatique (option PRCD), ainsi qu'un module d'Algorithmique Numérique en 1^{ière} année du cycle ingénieur. Chaque année ma charge d'enseignement correspond à 192 HETD auxquelles s'ajoutent environ 100 HETD supplémentaires. J'ai également une expérience d'enseignement à l'étranger ayant dispensé des cours au Gabon (niveau licence) en 2006 et des cours au Vietnam (niveau master, en anglais) depuis 2008.

J'étais également responsable pendant 6 ans de l'organisation des projets tuteurés de deuxième année de DUT et de licence professionnelle. Dans le cadre de nos licences professionnelles, les étudiants sont regroupés en équipe (2 à 5 étudiants par équipe) et doivent, en 4 semaines à temps plein, développer une application répondant aux besoins d'un client (le rôle du client étant assuré par un professionnel ou un enseignant). Pour la réalisation de certains projets, les étudiants peuvent utiliser du matériel multimédia (tablettes tactiles, smartphones, ...) et contribuer aux ressources documentaires associées (wiki). Ils doivent produire un poster, réaliser une démonstration auprès des clients et présenter leur travail devant un jury d'enseignants, lors d'une soutenance en anglais.

Entre 2007 et 2014, j'ai assuré la **direction des licences professionnelles** (Systèmes Informatiques et Logiciels) pour les spécialités ACPI (Assistant Chef de Projet Informatique) et DAWIN (Développeur en Applications Web et Images Numériques). Les tâches administratives correspondent :

- à la mise en place du plan d'enseignement et des emplois du temps,
- à la répartition des charges d'enseignement,
- au recrutement des étudiants et à la promotion de nos offres de formation,
- à la gestion des intervenants professionnels extérieurs au département,
- au suivi des étudiants, à la gestion des évaluations et à l'organisation des jurys,
- à la recherche et à l'affectation des stages,
- au pilotage du conseil de perfectionnement annuel.

Enfin, en collaboration avec E. Fleury, nous avons construit et présenté une **nouvelle offre de formation** (SSRL pour Spécialiste en Sécurité des Réseaux et Logiciels) au niveau licence professionnelle autour de la sécurité dans les réseaux et logiciels en nous appuyant sur une partie des contenus et de l'expérience du master CSI (Cryptologie et Sécurité Informatique). Cette proposition a été présentée, à la rentrée universitaire 2011, au CEVU de l'Université Bordeaux 1 qui a donné un avis favorable, avec un soutien fort des entreprises du secteur en Aquitaine. L'ouverture de cette formation a cependant été refusée par le ministère, en 2012 puis en 2013, et n'a pas été ré-étudiée depuis.

Mots-clés : **simulation numérique, parallélisme, calcul haute performance, recouvrement calcul/communication, irrégularité, hétérogénéité, ordonnancement, distribution de données, algèbre linéaire creuse.**

Ma problématique de recherche concerne les problèmes du calcul haute performance. Le thème général de mon travail concerne le calcul parallèle scientifique pour les problèmes réguliers et irréguliers et je m'intéresse en particulier aux problèmes de communication, de distribution et de prise en compte de l'hétérogénéité des plates-formes en permanente évolution.

► Dans mes travaux de thèse, j'ai présenté des contributions concernant la problématique du recouvrement calcul/communication.

J'ai proposé un modèle ainsi qu'un schéma de résolution permettant de déterminer analytiquement ou numériquement la taille optimale de paquets qui maximise le recouvrement dans les algorithmes macro-pipelines réguliers. Les résultats obtenus ont fait l'objet d'une présentation lors la conférence nationale *RenPar'8* [102] et internationale *Europar'96* [29]. La suite logique de ces travaux a été de proposer un calcul de la suite de tailles de paquets qui optimise le recouvrement pour le problème de la factorisation de Cholesky d'une matrice symétrique, la prise en compte de la symétrie constituant une première étape vers la prise en compte des problèmes irréguliers. Ces résultats ont fait l'objet d'une publication au congrès *RenPar'9* [101]. Des expérimentations ont été effectuées sur un Paragon Intel et un SP2 IBM pour valider cette étude sur des algorithmes classiques d'algèbre linéaire dense (factorisation *LU* et méthode *SOR*). Des gains significatifs ont été obtenus par l'utilisation de la bibliothèque *OPIUM* permettant le calcul de la taille optimale de paquets connaissant la complexité des calculs encadrant les communications et le modèle de communication.

Ce travail s'inscrit dans une démarche plus générale du recouvrement sur des architectures totalement hétérogènes et pour des applications irrégulières. J'ai débuté une étude de modélisation et la recherche des tailles de paquets qui minimisent le temps d'exécution pour un pipeline sur une plate-forme hétérogène. Enfin, dans le cas d'une architecture hétérogène ou hiérarchique (typiquement un cluster de nœuds multicœurs, on est amené à considérer le problème du recollement de pipelines présentant des grains de calcul différents qui m'avait conduit à développer une collaboration avec F. Silber-Chaussumier (Télécom et Management SudParis) en 2003.

► Ma principale contribution consiste à concevoir, développer et faire évoluer le solveur de la bibliothèque *PaStiX* qui fut le premier solveur direct à résoudre des systèmes avec plusieurs dizaines de millions d'inconnues, issus d'applications industrielles en trois dimensions [9].

Avec l'émergence des architectures à base de nœuds multicœurs (allant jusqu'à une centaine de cœurs par nœud), nous travaillons actuellement sur une version dédiée de *PaStiX* pour de telles plateformes de calcul. Ainsi nous nous sommes orientés vers une version de programmation hybride Thread/MPI du solveur qui a pour principe de partager les données d'un même nœud multicœurs via l'utilisation de threads POSIX. Outre l'avantage d'un système de communication plus performant, cette version offre des avantages algorithmiques très intéressants. *Cette version a permis de résoudre un problème avec 83 millions d'inconnues, en variable complexe, issu d'un code d'électromagnétisme 3D du CEA/CESTA, sur le supercalculateur TERA-10 du CEA (la résolution de ce système nécessite environ 5 PetaFlops et a été factorisé en 5 heures sur 768 processeurs). A notre connaissance, un système de cette taille et de ce type n'avait jamais été résolu précédemment par une méthode directe.*

Le solveur PaStiX est utilisé en production dans des codes du CEA/CESTA [11] et fait l'objet de collaborations avec EURATOM/CEA/CADARACHE autour de la fusion contrôlée et du réacteur ITER [6].

D'autre part, nous avons étendu les fonctionnalités du solveur pour pouvoir traiter des matrices non symétriques mais à structure symétrique ; dans ce cas, nous nous autorisons à faire du pivotage numérique statique ce qui implique d'appliquer, après résolution, une phase de raffinement itératif. Cette version de PaStiX est intégrée dans le code FluidBox développé dans le projet BACCHUS pour simuler les écoulements fluides instationnaires.

► Mes travaux de recherche concernent également l'étude de préconditionneurs basés sur une factorisation parallèle incomplète par blocs.

Le but de ces travaux est d'utiliser au mieux la technologie algorithmique parallèle par bloc qui est utilisée pour les solveurs directs haute performance pour développer des préconditionneurs parallèles de type Cholesky incomplet de manière adaptative et paramétrée, et qui soient robuste numériquement, suffisamment économes en mémoire et très performants en temps CPU (puissance tirée par processeur importante et bonne scalabilité).

Ces préconditionneurs sont alors intégrés dans des implémentations par blocs de méthodes de Krylov du type gradient conjugué ou GMRES. On recherche donc un compromis entre une diminution importante de la taille mémoire pour stocker les facteurs de la méthode directe, et une conservation d'une certaine dose de remplissage pour exploiter suffisamment les effets superscalaires dans les calculs BLAS3 du préconditionneur et atteindre globalement de bonnes performances en temps. Nous conservons les techniques de renumérotation et de distribution/ordonnancement du solveur direct pour avoir une implémentation parallèle efficace du calcul du préconditionneur et des itérés.

Un point crucial est donc le choix de la partition initiale pour la factorisation symbolique incomplète. Nous avons proposé récemment une méthode pour trouver des supernœuds approximatifs pour la factorisation $ILU(k)$. Les détails de cette démarche et l'analyse des premiers résultats de convergence pour des systèmes de grandes tailles ont été exposés lors de conférences SIAM sur les préconditionneurs et un article a été publié dans la revue *Parallel Computing* [8].

► Dans le cadre de la thèse d'Astrid Casadei, nous avons proposé des optimisations algorithmiques pour les principaux solveurs linéaires creux hybrides directs itératifs tels que HIPS, MaPHyS, PDSLIN ou ShyLU, qui sont basés sur une décomposition de domaine et une approche « complément de Schur ». Bien que ces solveurs soient moins coûteux en temps et en mémoire que leurs homologues directs, ils ne sont néanmoins pas exempts de surcoûts. Nous nous sommes intéressés à la question de l'équilibrage de la charge que pose la décomposition de domaine pour le calcul parallèle. Ce problème revient à partitionner le graphe d'adjacence de la matrice en autant de parties que de domaines désirés. Nous avons mis en évidence le fait que pour avoir un équilibrage correct des temps de calcul lors des phases les plus coûteuses d'un solveur hybride tel que MaPHyS, il faut à la fois équilibrer les domaines en termes de nombre de nœuds et de taille d'interface locale. Jusqu'à aujourd'hui, les partitionneurs de graphes tels que Scotch et MeTiS ne s'intéressaient toutefois qu'au premier critère (la taille des domaines) dans le contexte de la renumérotation des matrices creuses. Nous avons proposé plusieurs variantes des algorithmes existants afin de prendre également en compte l'équilibrage des interfaces locales et ces modifications ont été implémentées dans le partitionneur Scotch [14].

► Dans le cadre d'une initiative de recherche récente nous utilisons le moteur d'exécution générique StarPU (<http://starpupforge.inria.fr>) afin de proposer la première implémentation d'un

solveur creux direct sur des architectures hétérogènes multiCPUs et multiGPUs. Dans la thèse de Xavier Lacoste, nous avons étudié les bénéfices et les limites que peut nous apporter le remplacement de l'ordonnanceur natif, très spécialisé, du solveur PaStiX par deux systèmes d'exécution génériques : PaRSEC et StarPU. Pour cela, l'algorithme doit être décrit sous la forme d'un graphe de tâches qui est fourni aux systèmes d'exécution qui peuvent alors calculer une exécution optimisée de celui-ci pour maximiser l'efficacité de l'algorithme sur la machine de calcul visée. Une étude comparative des performances de PaStiX utilisant l'ordonnanceur interne, PaRSEC, et StarPU a été menée sur différentes architectures. L'analyse met en évidence les performances comparables des versions utilisant les systèmes d'exécution par rapport à l'ordonnanceur natif optimisé pour PaStiX. De plus ces implémentations permettent d'obtenir une accélération notable sur les machines hétérogènes en utilisant les accélérateurs tout en masquant la complexité de leur utilisation au développeur [15]. Afin de pouvoir utiliser ces travaux de manière efficace dans des codes parallèles de simulations, nous avons également présenté une interface distribuée, orientée éléments finis, permettant d'obtenir un assemblage optimisé de la matrice distribuée tout en masquant la complexité liée à la distribution des données à l'utilisateur.

Charges liées à la recherche

J'ai été rapporteur d'articles soumis aux revues : TOMS, JCP, PDCP et aux conférences : IPDPS, EuroPar, PMAA, RenPar.

J'étais co-organisateur de la conférence internationale "Preconditioning Techniques For Scientific And Industrial Applications" qui a lieu tous les deux ans et qui s'est déroulée à Bordeaux en mai 2011.

Je suis membre de l'organisation de la 9ième édition de la conférence "Parallel Matrix Algorithms and Applications" qui a lieu tous les deux ans et qui se déroulera à Bordeaux en juillet 2016.

Collaborations nationales

- **ANR CIS05 NUMASIS** : Dans le cadre de l'appel 2005 de l'ANR "Calcul intensif", j'ai co-encadré la thèse de Mathieu Faverge financée sur le projet NUMASIS. Nous nous sommes intéressés aux problèmes d'algorithmique parallèle et aux modèles de programmation sur architectures NUMA. Cela concerna aussi bien les algorithmes spécifiques au domaine applicatif de la sismique que les noyaux et bibliothèques de calcul classiques qui interviennent dans la conception des codes scientifiques de simulation. Ces résultats ont été présentés lors des conférences PARA'08 [20] et PMAA'08 [60].
- **ANR CIS06 ASTER** : Pour l'appel 2006 de l'ANR "Calcul intensif", avec Guido Huysmans, nous avons monté le projet ASTER dont je suis le responsable scientifique pour le partenaire Inria/LaBRI. Ce projet vise à développer et implémenter des méthodes pour améliorer les codes en Magnéto-Hydro-Dynamique afin de permettre la simulation des instabilités dans les plasmas. La compréhension de ces instabilités est très limitée et les pertes d'énergie induites sont un réel problème pour le réacteur ITER. Nous devons en particulier travailler sur la prise en compte de techniques de type raffinement de maillages adaptatifs et sur l'adaptation de nos solveurs. Les premiers résultats ont été présentés dans la revue *Plasma Physics and Controlled Fusion* [6]. Un article a été publié dans la revue *MagnetoHydroDynamics* [7].

- **ANR CIS06 SOLSTICE** : Mes travaux de recherche s’inscrivaient également dans le projet SOLSTICE de l’appel à projets 2006 de l’ANR “Calcul intensif”. L’objectif de ce projet était la conception et la mise en œuvre haute performance de solveurs linéaires parallèles efficaces pour résoudre des problèmes scientifiques complexes multi-physiques et multi-échelles de très grande taille, et leur intégration effective dans des codes applicatifs académiques et industriels dans le but de faire des simulations "grands challenges".
- **ANR COSINUS08 PETAL** : Mes travaux sur les factorisations incomplètes de type ILU s’inscrivent aussi dans le projet PETAL de l’appel à projets 2008 de l’ANR “Cosinus”. Il s’agit en particulier de comparer nos approches avec celles développées par d’autres membres du projet et qui présentent l’avantage d’être très peu consommatrices en ressource mémoire et calcul. Ce projet a été renouvelé en 2010 (PETALH). En particulier, une méthode a été développée pour réduire les surcoûts mémoires générés lors de la construction du complément de Schur dans un solveur direct, ces résultats seront présentés à la conférence SIAM Linear Algebra 2012 [54].
- **ANR MN11 ANEMOS** : Il s’agit de la suite du projet ASTER dans laquelle nous nous engageons, avec nos partenaires, à faire des progrès significatifs dans la compréhension de la physique inhérente aux stratégies de contrôle actif des instabilités MHD du bord du plasma dans les Tokamaks (et ITER en particulier). Ceci nécessite de poursuivre nos efforts dans l’optimisation des composantes logicielles et des solveurs en particulier. Nous avons ainsi pu mettre en œuvre une nouvelle interface entre nos solveurs et le code numérique JOEREK afin de minimiser les surcoûts mémoires et les redistributions de données. Enfin, nous avons commencé une étude pour adapter un solveur direct creux aux architectures multicœurs/multigpus. Les premiers résultats ont été présentés au workshop HCW de la conférence IPDPS 2014 [15].
- **ANR MN13 SOLHAR** : Ce projet a pour but d’étudier et de mettre au point des algorithmes et des modèles de programmation parallèles pour implanter des méthodes directes de résolution de systèmes linéaires creux sur les plates-formes émergentes équipées d’accélérateurs. Le but à long terme de ce projet est de développer une solution logicielle fournissant un solveur basé sur les méthodes directes pour résoudre des systèmes creux d’équations linéaires. A ce jour, les approches proposées pour atteindre cet objectif sont essentiellement basées sur un simple modèle de délégation de certaines tâches de calcul aux accélérateurs et requièrent une optimisation manuelle du code. Au contraire, l’approche proposée s’articule autour de trois axes de recherche : algèbre linéaire, moteurs d’exécution, ordonnancement.

Collaborations internationales

- **Collaboration NSF/Inria** : Nos travaux ont permis d’établir des échanges internationaux dans le cadre d’un contrat NSF/Inria et lors de deux séjours à l’université de Minneapolis, sur invitation de Y. Saad. J’ai participé à l’équipe associée Phyleas, portée par J. Roman, impliquant Inria et l’université du Minnesota.
- **Collaboration avec le JAEA** : En 2007, j’ai participé au lancement d’un projet de collaboration avec le JAEA (Japan Atomic Energy Agency, qui participe au projet ITER), initié par M. Daydé, professeur à l’INPT. Le JAEA a accès à l’infrastructure du projet NAREGI (NAtional REsearch Grid Initiative), un équivalent japonais du projet Grid’5000 en France, et développe également ses propres solveurs en algèbre linéaire pour ses applications de simulation

numérique. Cette collaboration a donné lieu à plusieurs présentations [18, 19, 79].

- **Collaboration avec l'ICL** : Depuis 2011, dans le cadre de l'équipe associée MORSE, j'ai mis en place une collaboration avec l'équipe ICL (Innovative Computing Laboratory) dirigée par Jack Dongarra. Nous avons ainsi utilisé l'ordonnanceur générique PaRSEC (<http://icl.eecs.utk.edu/dague>), développé par l'équipe de G. Bosilca, afin de proposer la première implémentation d'un solveur creux direct sur des architectures multiCPUs et multiGPUs. Les premiers résultats ont été présentés au workshop HCW de la conférence IPDPS 2014 [15].
- **Collaboration avec le LNCC** : Depuis 2012, je participe aux échanges organisés dans le cadre du projet HOSCAR avec différents organismes brésiliens et en particulier le LNCC. Cette collaboration sera poursuivie avec le projet HPC4E du l'appel Europe-Brésil H2020.
- **Collaboration avec Stanford** : Depuis 2013, dans le cadre de l'équipe associée FAST-LA, j'ai mis en place une collaboration avec l'équipe d'Eric Darve sur l'application des techniques de compression utilisant des matrices hiérarchiques pour l'accélération et la réduction de la complexité des solveurs directes. Une *bourse Chateaubrilland* a été accordée en 2014, dans le cadre de cette collaboration, à Clif Robert Dudley, étudiant de master à Stanford. Cet échange a permis de définir un premier prototype utilisé comme base pour le nouveau solveur développé par Grégoire Pichon dans sa thèse débutée en octobre 2015 [41].

Valorisations et transferts industriels

- **Contrat de valorisation avec le CEA/CESTA** : Dans le cadre de mes travaux de recherche, j'ai participé à des contrats [111, 109, 108, 107] avec obligation de résultat passés avec le CEA/CESTA. Il s'agissait d'études d'applicabilité de nos méthodes aux problèmes étudiés par l'équipe "Analyse Numérique" du CEA et de transférer nos développements logiciels. J'ai également travaillé sur la parallélisation du code MIRO, un code qui simule la propagation d'un faisceau laser dans une chaîne de composants optiques pour le laser Mégajoule [110]. J'ai pu valoriser mes compétences dans le domaine des techniques de parallélisation dans le cadre de plusieurs contrats d'expertise passés avec le CEA. Depuis septembre 2015, je suis **conseiller scientifique au CEA CESTA** avec une expertise pour le domaine du HPC.
- **Transfert industriel avec EURATOM/CEA/CADARACHE** : Depuis 2006, j'ai assuré le transfert et le suivi du solveur PaStiX dans le code de simulation JOEREK qui a pour objectif d'améliorer la compréhension des instabilités au bord des plasmas, en particulier pour le contrôle du réacteur ITER utilisant le principe de la fusion thermonucléaire. Cette collaboration nous a conduits à améliorer l'efficacité de l'interface proposée entre le solveur et le code de simulation, tant du point de vue des performances que de la consommation mémoire.
- **Contrats de collaboration avec EDF** : En 2006, j'ai encadré le travail de Guilhem Caramel sur "l'optimisation des performances des outils de calcul de neutronique des cœurs" dans le cadre d'un contrat [106] passé avec EDF. Cette étude s'est prolongée dans le cadre de la thèse de Bruno Lathuilière et du contrat de collaboration passé avec EDF-R&D. J'ai co-encadré ce travail portant sur la parallélisation d'une méthode de décomposition de domaine des équations SPN de la neutronique pour les études d'accident. Les résultats ont été présentés lors des conférences CSE'08 [17], IMACS'08 [62], PMAA'08 [61] et *Reactor Physics* [58]. Un article a été publié dans la revue *Journal of Computational Physics* [5]. D'autre part, les calculs de

transport par la méthode SN en neutronique forment un domaine de recherche où il existe une compétition internationale importante. Ces codes sont très consommateurs en temps de calcul et, depuis l'origine, ils sont utilisés comme benchmark pour évaluer les architectures HPC dont ils approchent les performances crêtes. Dans le cadre d'un contrat de collaboration passé avec EDF-R&D, l'objectif de la thèse de Salli Moustafa était la mise au point d'une méthode de résolution massivement parallèle de l'équation de transport neutronique. La réduction du temps de retour à moins d'une heure, pour des calculs impliquant plus de 10 milliards d'inconnues, constitue une avancée importante pour la simulation des coeurs des centrales nucléaires. Pour y parvenir, il est nécessaire d'utiliser les techniques de parallélisation les plus avancées impliquant plusieurs algorithmes parallèles imbriqués ainsi que plusieurs paradigmes de programmations parallèles (passage de message, multi-threading et instructions SIMD). Une stratégie algorithmique adaptée au cas particulier des systèmes diffusifs allié à un très haut niveau de parallélisme a conduit à la réalisation d'un outil unique pour la simulation des centrales nucléaires [3, 13]. Un article est en préparation pour une soumission dans la revue *Journal of Computational Physics*.

- **Contrats d'expertise avec Algo'Tech** : En 2014, avec le soutien de l'Initiative HPC-PME (lancée par Bpifrance, GENCI et Inria), la société Algo'Tech, basée à Bidart, spécialisée dans l'édition de logiciels de schématique électrique et de câblage pour les travaux publics, les machines industrielles et les systèmes embarqués, a réalisé avec succès le passage au HPC (High Performance Computing ou calcul haute performance) de son logiciel de câblage électrique. La PME a bénéficié d'un transfert de compétences dans le cadre de travaux sur le modèle du thésard-conseil réalisé par le doctorant Xavier Lacoste [103]. Suite à cette expertise, nous avons obtenu des supports européens avec les appels FORTISSIMO en 2014 et PRACE-4IP en 2015. L'utilisation du solveur PaStiX a ainsi permis de réaliser le saut technologique nécessaire au développement d'une version de son logiciel adaptée au HPC.

Logiciels

Je suis l'architecte et j'assure le développement ainsi que le suivi de la bibliothèque PaStiX [118] qui implémente un solveur haute performance pour la résolution de grands systèmes linéaires basé sur une approche directe supernodale.

Un site web donnant l'accès au téléchargement et à la documentation de l'interface est disponible sur : <http://pastix.gforge.inria.fr>. Ce logiciel est maintenant placé sous licence libre CeCILL-C et la première distribution publique date de septembre 2006 sur la forge Inria. Cette distribution, qui représente actuellement plus de 120000 lignes de code, est constamment mise à jour pour tenir compte des améliorations algorithmiques apportées par mes travaux, ce processus étant matérialisé par la mise à disposition de la communauté de versions successives avec 2 révisions majeures par an. La distribution logicielle PaStiX est très largement diffusée et utilisée, et a fait l'objet de plus de 60000 téléchargements (dans le TOP20 des logiciels développés sur la forge Inria).

Ce solveur est reconnu par la communauté pour être un des plus performants sur les architectures récentes constituées de nœuds multicœurs. Très récemment, nous avons pu proposer la première implémentation d'un solveur creux direct sur des architectures hétérogènes disposant d'accélérateurs de type GPU.

Ce solveur est intégré dans plusieurs autres logiciels, qui contribuent à sa diffusion indirecte, par exemple :

- PETSc (<http://www.mcs.anl.gov/petsc>) : une bibliothèque de fonctions en C permettant de gérer des vecteurs et des matrices creuses et de résoudre les systèmes linéaires correspondants avec des solveurs directs ou itératifs ;
- Eigen (<http://eigen.tuxfamily.org>) : une bibliothèque template en C++ d'algèbre linéaire.

Enfin, parmi les principaux utilisateurs, nous pouvons citer ceux avec qui nous entretenons des collaborations scientifiques régulières :

- CEA/CESTA : pour des codes en production de mécanique des structures et d'électromagnétisme ;
- EURATOM/CEA/CADARACHE : pour un code de Magnéto-Hydro-Dynamique dont le but est de contribuer au dimensionnement et au contrôle du réacteur ITER.
- Algo'Tech : dans le cadre de l'initiative HPC PME, la société Algo'Tech utilise le solveur PaStiX dans ses codes de simulation de systèmes électriques et électromagnétiques pour ses clients issus aussi bien de l'aéronautique que de l'automobile.

Encadrement d'étudiants dans le cadre de travaux de recherche

Entre 2006 et 2015, j'ai co-encadré cinq thèses :

- la thèse de Mathieu Faverge [112] financée sur le projet NUMASIS qui a soutenu sa thèse en décembre 2009. Les nouvelles architectures de calcul intensif intègrent de plus en plus de microprocesseurs qui eux-mêmes intègrent un nombre croissant de cœurs de calcul. Cette multiplication des unités de calcul dans les architectures a fait apparaître des topologies fortement hiérarchiques. Ces architectures sont dites NUMA. Nous avons étudié un ordonnancement dynamique adapté aux architectures NUMA pour un solveur direct creux supernodal. Les structures de données du solveur, ainsi que les schémas de communication ont dû être modifiés pour s'adapter aux caractéristiques de ces architectures et à un ordonnancement dynamique. Nous nous sommes également intéressés à l'adaptation dynamique du grain de calcul pour exploiter au mieux les architectures multi-cœurs et la mémoire partagée ;
- la thèse de Bruno Lathuilière [113] sur un contrat de collaboration passé avec EDF-R&D qui a soutenu sa thèse en janvier 2010. Les calculs de réactivité constituent une brique fondamentale dans la simulation des coeurs des réacteurs nucléaires. Ceux-ci conduisent à la résolution de problèmes aux valeurs propres généralisées via l'algorithme de la puissance inverse. A chaque itération, on est amené à résoudre un système linéaire de manière approchée via un algorithme d'itérations imbriquées. Au cours de cette thèse, nous avons étudié une méthode de décomposition de domaine de type Schur dual. Plusieurs placements de l'algorithme de décomposition de domaine au sein du système d'itérations imbriquées ont été considérés. Les résultats obtenus permettent d'envisager l'industrialisation de la méthodologie associée ;
- la thèse de Salli Moustafa [116] sur un contrat de collaboration passé avec EDF-R&D qui a soutenu sa thèse en décembre 2015. L'objectif de ce travail est la réalisation d'un solveur de transport neutronique massivement parallèle conduisant à la simulation 3D en transport exact d'un coeur REP discrétisé au niveau du crayon, pour des quadratures angulaires comportant plusieurs centaines de directions et pour quelques dizaines de groupes d'énergie. L'utilisation

du parallélisme massif va permettre de réaliser des calculs cinétiques en transport exact et de les comparer aux calculs de COCAGNE (solveur de transport simplifié SPN). On devrait alors obtenir pour la première fois une estimation fiable de la précision des différents niveaux d'approximation utilisés pour la simulation 3D de la dynamique rapide des coeurs de centrale.

- la thèse de Xavier Lacoste [114], financée sur le projet ANEMOS. [VOIR PROJET DE RECHERCHE]
- la thèse d'Astrid Casadei [115] financée sur une allocation de recherche ministérielle. [VOIR PROJET DE RECHERCHE]

Depuis la rentrée 2015, **j'encadre la thèse de Grégoire Pichon**, financée sur un contrat DGA. [VOIR PROJET DE RECHERCHE]

L'objectif final de cette thèse est de gagner un ordre de grandeur sur la taille des problèmes que l'on sait traiter aujourd'hui avec la version du code PaStiX qui exploite déjà les technologies informatiques les plus innovantes (MPI+threads, supports d'exécution à base de tâches).

Dans un premier temps, l'étude se focalisera sur la compression du complément de Schur, particulièrement consommateur en mémoire et central dans le cadre des solveurs hybrides. Ce travail impactera naturellement les travaux sur le solveur hybride MaPHyS et sera conduit en synergie avec les actions en cours autour de ce code.

Dans un deuxième temps, nous développerons un prototype du solveur PaStiX exploitant les techniques de compression des supernoeuds sur plusieurs niveaux de l'arbre d'élimination. Ceci réduira l'empreinte mémoire ; par ailleurs, la réduction du nombre d'opérations lors de la factorisation numérique et des étapes de descente remontée pour la résolution sera étudiée. Il faudra alors développer et analyser de nouvelles techniques de renumérotation pour préserver les arbres de compression et ainsi obtenir un couplage efficace et compatible entre l'arbre de compression des matrices hiérarchiques et les techniques de numérotation à base de dissections emboîtées qui cherchent à minimiser le remplissage. Compte tenu de la réduction du coût mémoire lié au remplissage, d'autres métriques pourront également être considérées pour guider la numérotation.

Ce travail s'appuiera sur les outils développés par l'équipe d'Eric Darve (Stanford University) utilisant des matrices hiérarchiques pour la résolution de systèmes linéaires denses (bbFMM, FLIPACK, HOLDR_Solver). Il est à noter que le nouveau paradigme de programmation exploité dans le solveur PaStiX sera une plate-forme facilitant l'expérimentation modulaire de ces noyaux de calcul.

En plus de l'encadrement des quatre thèses citées précédemment, j'ai également eu l'occasion d'encadrer le travail de nombreux étudiants dans le cadre de mes travaux de recherche. En particulier, j'ai encadré le postdoctorat de H. Sellama financé sur le projet ASTER. Le travail consistait à intégrer une procédure de raffinement dans le code JOREK, tout d'abord au début de la simulation (pour l'équilibre) puis en apportant des modifications pour obtenir un raffinement adaptatif (au cours de la simulation). Les étapes de raffinement et de dé-raffinement des éléments finis de type Bézier ont été implémentées pour le calcul à l'équilibre. Le raffinement adaptatif (au cours de la simulation) est également disponible et les premières simulations illustrant l'intérêt de l'approche ont été menées comme par exemple l'injection de "glaçons" pour le contrôle de la densité du plasma. Enfin, le traitement particulier de la région du point-X dans géométrie des Tokamaks était une difficulté qui a été prise en compte lors du raffinement. Un article en revue est en cours de rédaction sur ce travail.

Formations avancées et autres activités

Dans le cadre d'écoles jeunes-chercheurs soutenues par le CNRS (ICARE'97, HiPerf'2000 et Grid'2002), rassemblant plus d'une centaine de participants, j'ai eu l'occasion de présenter mes travaux de recherche. De plus, j'ai pris part à l'organisation de plusieurs formations avancées dans le domaine du parallélisme et de ses applications :

- une formation généraliste dispensée à des personnels du CEA en juin 1996 ;
- une formation nationale sur la programmation avancée des calculateurs parallèles à mémoire distribuée (MPI et HPF), organisée par le CINES et le LaBRI en mars et mai 1997 ;
- une formation, de même contenu (MPI et HPF), organisée au LaBRI et dispensée à des personnels du CEA en février et mars 1998.

J'ai été invité à présenter mes travaux lors de :

- l'école CEA-EDF-Inria sur le calcul scientifique intensif à Rocquencourt en 2006 [98] ;
- la formation sur l'informatique scientifique pour le calcul organisée par le CNRS à Sète en 2008 [97] ;
- la formation sur les solveurs de systèmes linéaires de grande taille organisée par le CNRS à Lyon en 2010 [95] ;
- l'école d'été du CEMRACS sur les modèles numériques pour la Fusion à Marseille en 2010 [96] ;
- l'école de mécanique des fluides numériques à Roscoff en 2011 [94] ;
- la formation sur l'algèbre linéaire creuse parallèle organisée par la Maison de la Simulation, chaque année, depuis 2011 [84, 88, 90, 93] ;
- l'école d'été du CEMRACS sur les méthodes numériques et algorithmes pour architectures pétaflopiques à Marseille en 2012 [92] ;
- l'école thématique du CEA sur la simulation numérique à Fréjus en 2013 [91] ;
- l'école internationale ITER "High Performance Computing in Fusion Science" à Aix-en-Provence en 2014 [87] ;
- la formation du CNRS dédiée au problème de Poisson à Paris en 2015 [86].

En collaboration avec l'équipe système du LaBRI, j'ai administré le calculateur parallèle du laboratoire (IBM SP2 puis SP3). Les moyens de calcul sont maintenant gérés au niveau régional avec le partage des ressources du MCIA (Mésocentre de Calcul Intensif Aquitain) et de PlaFRIM (Plateforme Fédérative pour la Recherche en Informatique et Mathématiques). Depuis la restructuration du mésocentre en 2010, je représente le LaBRI au comité scientifique et je participe au bureau opérationnel qui anime au quotidien les différentes missions.

Enfin, depuis 10 ans, je suis le correspondant, pour les projets BACCHUS et HiePACS, des demandes de ressources informatiques sur les centres de calcul nationaux CINES, IDRIS et CCRT maintenant regroupés dans le cadre d'un appel commun sous la direction du GENCI (Grand Equipement National de Calcul Intensif). Depuis 2008, je suis expert pour le comité thématique 6 (Maths-Info) du GENCI avec la mission d'évaluer et classer les demandes de ressources sur les calculateurs des centres nationaux.

Du fait des quantités considérables de données et de calculs mises en jeu, la résolution des simulations numériques en vraie grandeur imaginées aujourd’hui par les chercheurs n’est réalisable qu’en ayant recours efficacement au calcul parallèle haute performance au sens large. Dans la plupart des simulations numériques, l’étape la plus consommatrice en ressources (CPU et mémoire) est la résolution de systèmes linéaires creux de très grande taille. En plus de la difficulté numérique inhérente à certains modèles physiques employés, l’augmentation de la taille des systèmes linéaires conduit même les méthodes réputées scalables, telles que les méthodes itératives classiques, à ne plus converger de manière efficace.

Le solveur PaStiX intègre une gestion fine et efficace des multiples niveaux de parallélisme exploitables dans un solveur linéaire creux. En effet, les nouvelles architectures de calcul intensif intègrent de plus en plus de microprocesseurs qui eux-mêmes intègrent un nombre croissant de cœurs de calcul. Cette multiplication des unités de calcul dans les architectures a fait apparaître des topologies fortement hiérarchiques (NUMA). Nous avons donc travaillé sur les problèmes d’ordonnancement dynamique adaptés à ces architectures. Nous nous intéresserons également à l’adaptation dynamique du grain de calcul pour exploiter au mieux les architectures multicœurs et la mémoire partagée. L’écriture générique du schéma de résolution, sous la forme d’un graphe de tâches, permet de disposer d’une plateforme efficace pour concevoir et développer des ordonnancements pour des architectures encore plus hétérogènes, comme celles qui mixent CPUs et GPGPUs.

Le solveur HIPS s’intéresse à des méthodes hybrides directes/itératives, qui utilisent une approche de type « complément de Schur » basée sur une décomposition de domaines algébrique : les inconnues associées à l’intérieur des sous-domaines sont éliminées en utilisant un solveur direct creux ce qui permet de se ramener à un système réduit défini uniquement sur l’interface. Ce système de plus petite taille est résolu grâce à une méthode itérative préconditionnée par une factorisation incomplète. Un réordonnancement adapté de l’interface permet la construction d’un préconditionneur global et naturellement parallèle du complément de Schur, compatible avec la décomposition de domaines. Cette méthode procure un fort degré de parallélisme car elle est basée sur une décomposition de domaines tout en procurant une bonne robustesse par rapport à des approches plus classiques telles que les méthodes « Schwarz additive » et les factorisations ILU classiques.

Cependant, l’arrivée des architectures exa-scale, de type Blue Waters (celle-ci comprenant 150.000 cœurs), incite à anticiper de nouveaux challenges. Par exemple le partitionneur de graphes et de maillages Scotch a déjà permis de traiter des cas à 2 milliards d’inconnues. L’étape limitante pour atteindre ce nombre d’inconnues dans les simulations numériques est maintenant clairement la résolution des systèmes linéaires. Il est donc critique de travailler sur la conception et la réalisation d’un solveur linéaire creux capable d’exploiter les nouvelles architectures de ces machines exa-scale. Je propose donc de développer mes activités de recherche en rejoignant l’équipe HiePACS, pour contribuer aux domaines suivants avec des apports à la fois méthodologiques et algorithmiques.

► Prétraitements numériques et partitionnement parallèle.

Une première étape consiste à définir des fonctionnalités pour la prise en compte de critères numériques dans le choix du partitionnement pour nos différentes classes de solveurs. En effet, les critères de partitionnement sont basés sur des informations topologiques dans le cas des maillages ou des mé-

triques de structure dans le cas des graphes. Dans le cas du logiciel Scotch, il est nécessaire de prendre en compte des poids sur les arêtes et, par exemple, de modifier la contraction du graphe dans l'approche multi-niveaux. Ce travail fait partie des objectifs fixés dans l'ANR PETALH et nous souhaitons recruter un candidat en postdoctorat à la prochaine rentrée universitaire.

Dans la cas d'une factorisation LU, les techniques de pivotage partiel ont pour inconvénients d'augmenter à la fois le temps d'exécution et le remplissage de la matrice. Il est donc pertinent d'essayer de prétraiter numériquement la matrice avant la factorisation soit pour réduire le nombre d'inconnues à pivoter soit pour renforcer la stabilité numérique lorsqu'on utilise des techniques de pivotage statique. Afin de stabiliser la convergence dans le cadre des méthodes hybrides et la robustesse des solveurs directs basés sur un pivotage statique, il est intéressant de poursuivre la conception et l'implémentation d'algorithmes permettant de construire une permutation non-symétrique correspondant à une renumérotation des inconnues. Les algorithmes de recherche d'un couplage de poids maximal dans un graphe biparti sont connus pour donner de bons résultats mais restent coûteux en complexité et difficiles à paralléliser. L'objectif de cet axe de recherche est d'obtenir une méthode parallèle pour maximiser la trace d'une matrice tout en maximisant le coefficient de symétrie du graphe associé. Des premiers résultats ont été obtenus dans le cadre de la thèse de S. Pralet au CERFACS et constituent une base de départ solide. Ces travaux intéressent en particulier l'équipe Inria ROMA avec les activités de B. Uçar et l'équipe du RAL (Rutherford Appleton Laboratory) avec J.A. Scott, et devront donc être menés de manière concertée.

► Solveur hybride par décomposition de domaines algébrique.

Dans le contexte des méthodes hybrides basées sur une approche par complément de Schur, il apparaît critique de contrôler l'équilibrage de la taille des interfaces. Une piste prometteuse consiste à prendre en compte la taille réelle des interfaces dans l'étape récursive d'une méthode de dissection emboîtée. Cette nouvelle activité de recherche est menée en collaboration avec Y. Saad de l'Université du Minnesota et E. Ng dans le cadre de l'équipe associée FastLA entre l'équipe HiePACS, le Lawrence Berkeley National Laboratory et l'université de Stanford. Trouver de nouveaux algorithmes et les implémenter de manière générique dans les principaux outils de partitionnement de graphe constitueront un défi et une avancée majeure pour la communauté des chercheurs qui travaillent sur les approches hybrides pour la résolution de très grands systèmes linéaires creux hors de portée des méthodes directes.

D'un point de vue algorithmique, il est également important de concevoir un couplage efficace des méthodes directes et itératives. Un tel couplage doit permettre une gestion optimale de tous les niveaux de parallélisme. Les approches hybrides, basées sur une décomposition de domaine algébrique utilisant une approche de type complément de Schur, sont actuellement très largement étudiées. Dans cette méthode, un solveur direct est utilisé comme une brique de base pour chaque sous-domaine de la matrice, mais les surcoûts mémoires, même s'ils sont réduits par rapport à une approche directe complète, restent un facteur limitant. Il est ainsi important de concevoir et étudier des méthodes pour réduire les surcoûts mémoires générés lors de la construction du complément de Schur dans un solveur direct. De plus, à l'issue d'une étude comparative, il sera nécessaire de faire converger les deux solveurs hybrides HIPS et MaPHYS pour aller vers une solution logicielle consolidée et étendue. Ce travail prendra naturellement sa place dans l'ADT MaPHYS@exa de l'Action d'Envergure C2S@exa coordonnée par S. Lanteri, dans laquelle je suis impliqué au titre du pôle "Numerical Linear Algebra".

Les travaux correspondant à ces deux derniers points ont été abordés durant la thèse d'A. Casadei

et ils laissent clairement entrevoir des perspectives scientifiques intéressantes.

Un autre type de couplage mixant des méthodes directes et multigrilles a été regardé dans les travaux de thèse de M. Chanaud. Dans un souci de robustesse, une idée originale est de conserver un solveur direct parallèle pour effectuer les résolutions sur le niveau grossier de la méthode multigrille. L'application d'une méthode directe sur le maillage grossier constitue le point d'entrée de l'algorithme multigrille et permet également de distribuer le système. En particulier, la distribution des données issues du solveur direct parallèle PaStiX est utilisée pour piloter la définition et le traitement des niveaux plus fins. Cette étude réalisée pour la résolution des équations de Maxwell tridimensionnelles mérite d'être poursuivie et étendue à d'autres applications dans le cadre du nouveau partenariat entre Inria et le CEA/CESTA.

► Ordonnancements pour architectures hétérogènes.

Il s'agit d'adapter des algorithmes d'ordonnement pour tirer parti d'un environnement d'exécution exploitant les architectures mixte CPUs/GPGPUs. Nous devons étudier la possibilité de remplacer les ordonnanceurs dynamiques basés sur des stratégies internes par des supports d'exécution génériques. Sur des noeuds de calcul disposant d'accélérateurs, ces supports d'exécution offrent la possibilité de dérouler le graphe de tâches de la factorisation numérique en s'exonérant la prise en compte de la complexité du matériel qui est prise en charge par le moteur d'exécution qui analyse le graphe et ses dépendances. Les premiers résultats montrent qu'une approche basée sur un DAG (Directed Acyclic Graph) offre une approche de programmation uniforme et portable pour réaliser du calcul haute performance sur des problèmes irréguliers d'algèbre linéaire creuse sur des nœuds hétérogènes.

Dans le cadre de l'équipe associée MORSE, nous avons ainsi pu évaluer le moteur d'exécution générique StarPU (<http://starpu.gforge.inria.fr>) en collaboration avec l'équipe RUNTIME Inria. Dans le cadre de la collaboration avec l'équipe ICL (Université du Tennessee) nous utilisons également le support d'exécution générique ParSEC (<http://icl.eecs.utk.edu/dague>) afin de proposer la première implémentation d'un solveur creux direct sur des architectures hétérogènes multiCPUs et multiGPUs.

De nombreux travaux restent cependant à mener. Dans le cas d'une approche supernodale, les structures de données peuvent être compactées de manière à optimiser l'efficacité des noyaux de calcul. En particulier, l'opération de produit matriciel a été particulièrement étudiée, mais il est nécessaire d'améliorer les performances de l'opération d'ajout matriciel lors des mises à jour dans les structures creuses des supernœuds. De plus, il sera nécessaire de fusionner les sous-arbres en bas de l'arbre d'élimination pour augmenter la granularité des tâches et ainsi réduire les surcoûts liés à l'ordonnement. Il est également possible de regrouper certaines branches de l'arbre comme dans les travaux regardés par A. Hugo dans sa thèse. Une autre manière de contrôler la granularité est d'utiliser les techniques d'amalgamation que nous avons proposées [8] pour alimenter à la fois le parallélisme potentiel sur les multiples cœurs de calcul et la réutilisabilité des données sur les accélérateurs de type GPU. Concernant une mise en œuvre sur des architectures à mémoires distribuées, il s'agira de trouver une manière de guider l'ordonnement des communications et de définir des fonctions de réduction qui seront proposées dans les futures versions des supports d'exécution génériques.

Les travaux correspondant à cette section ont été regardés dans la thèse de X. Lacoste et ils laissent clairement entrevoir des perspectives scientifiques intéressantes.

Un autre aspect critique du problème d'ordonnancement concerne la gestion optimale des ressources mémoires. Une variante de l'algorithme de "proportional mapping" sous contraintes mémoires dans le contexte d'un solveur multifrontal est actuellement regardée par l'équipe MUMPS. La méthode supernodale fan-in propose naturellement un compromis efficacité/mémoire en permettant de réduire les pics mémoire en transférant les contributions directement sur la cible "finale" tout en limitant le volume de communication ; la contrepartie étant d'engendrer un schéma de communication plus complexe et moins régulier. Il semble toutefois pertinent de guider l'étape de distribution des données en utilisant une borne supérieure des ressources mémoires pour les sous-arbres de l'arbre d'élimination supernodal. D'une manière connexe, les solutions proposées par E. Agullo dans sa thèse peuvent être adaptées aux arbres d'élimination des méthodes supernodales.

► \mathcal{H} -Matrix.

Au cœur d'un grand nombre de ces outils de simulations se trouvent des résolutions de systèmes linéaires de grandes tailles qui représentent souvent la partie dominante du temps de calcul. Ces solveurs linéaires s'appuient sur une grande variété de méthodes numériques et algorithmes. Numériquement et informatiquement cela se traduit par une hiérarchie de méthodes dont la base est le calcul matriciel dense, celui-ci est exploité par les méthodes de factorisation creuses, ces dernières sont une des composantes cruciales des méthodes hybrides qui combinent des approches directes et itératives.

De la communauté équations intégrales et éléments frontières a émergé, ces dernières années, des techniques de compression pour matrices denses globalement regroupées sous le vocable de \mathcal{H} -matrices. Ces techniques reposent sur une représentation "data sparse" des opérateurs linéaires et exploitent l'information géométrique du maillage sous-jacent pour définir le "clustering", i.e., la structure hiérarchique de la matrice dense. Nous proposons de prolonger ces idées dans le cadre purement algébrique des méthodes de factorisation de matrices creuses en substituant le calcul matriciel dense "classique" par du calcul \mathcal{H} -matriciel afin de réduire l'empreinte mémoire des solveurs directs creux qui constitue le principal frein au passage à l'échelle de ces approches pour des simulations 3D.

On souhaite étudier l'impact numérique des techniques de compression exploitant l'arithmétique des matrices hiérarchiques (\mathcal{H} -matrix) dans les blocs denses qui apparaissent lors de la factorisation numérique. L'utilisation des familles de matrices hiérarchiques repose sur une approximation de rang faible des blocs denses dont la structure est définie récursivement.

Quelques tentatives ont récemment été réalisées. X. S. Li (Lawrence Berkeley Nat. Lab.) a étudié l'utilisation des matrices HSS dans le cadre des factorisations de matrices creuses par méthode directe multi-frontale pour la résolution d'équations 3D pour des problèmes de diffusion et d'Helmholtz. Les bibliothèques MUMPS et CHOLMOD ont également évalué l'utilisation d'approximations de rang faible avec des approches non-hiérarchiques.

Ce travail s'inscrit dans une stratégie scientifique globale et fait le lien naturel entre plusieurs initiatives qui ont débuté récemment impliquant les chercheurs de l'équipe HiePACS et ses partenaires industriels ; deux de ces actions font l'objet de financements DGA.

Dans le projet DGA Rapid Hi-Box (IMACS, Airbus Group Innovation et Inria), qui vise à développer une boîte à outils pour des simulations parallèles en BEM, la bibliothèque parallèle de calcul \mathcal{H} -matrice d'Airbus est en cours de consolidation. Dans ce projet, la construction de la structure hiérarchique des matrices est complètement bâtie à partir de l'information géométrique du maillage

surfacique. Dans le cadre de la thèse de Grégoire Pichon, l'objectif est la résolution de systèmes linéaires creux issus de formulations FEM ; la définition de la structure hiérarchique des matrices denses qui apparaissent lors de la factorisation sera purement algébrique.

Par ailleurs, le résultat de cette étude conduira à une nouvelle version d'un solveur creux parallèle qui pourra être intégré dans les solveurs hybrides développés dans HiePACS. Ces solveurs hybrides sont en particulier utilisés dans le cadre du projet DGA TECSER ; ce dernier projet est en synergie avec le travail postdoctoral de Yuval Harness, qui s'intéresse à l'impact sur la mémoire et sur les aspects numériques de l'application de calcul \mathcal{H} -matriciel dans le solveur hybride MaPHyS. Ces travaux sont en particulier menés dans le cadre de l'équipe associée FastLA avec l'Université de Stanford et le Lawrence Berkeley Nat. Lab. (financement Inria pour des missions/visites afin de consolider des actions de recherche avec des équipes étrangères internationalement reconnues, celles de E. Darve et X.S. Li dans notre cas).

Le premier jalon sera de générer le complément de Schur ou les premiers niveaux de l'arbre d'élimination de manière à évaluer les taux de compression envisageables suivant les différents types de matrices hiérarchiques et la précision de l'approximation. Il sera alors possible de déterminer des bornes pour les gains en mémoire et en nombre d'opérations et ainsi connaître les tailles de bloc à partir desquelles la compression sera profitable en pratique. En particulier, il sera nécessaire de déterminer l'efficacité des algorithmes existants pour extraire les approximations de rang faible.

Dans un second jalon, un prototype du solveur sera développé dans lequel la compression sera appliquée sur les supernoeuds les plus larges. Néanmoins, pour simplifier la validation numérique, le solveur se basera sur un stockage non compressé des données mais utilisera les approximations hiérarchiques pour les opérations. Ainsi on pourra conserver les structures de données natives du solveur tout en utilisant les outils de compression hiérarchiques. Bien que non optimal, ce prototype permettra d'estimer l'accélération potentielle du solveur et de sa stabilité numérique.

Enfin, outre l'implémentation complète de tous les noyaux de calcul avec leurs structures de données hiérarchiques, nous considérerons plusieurs types de compression et mènerons une étude de complexité sur la chaîne complète du solveur suivant la classe de matrice hiérarchique considérée (strong/weak \mathcal{H} -matrix ou \mathcal{H}^2 -matrix) et la classe du problème visé (e.g. problème de diffusion/élasticité ou propagation d'ondes de type Helmholtz).

La validation et l'évaluation des performances de ce nouveau solveur seront réalisées sur des cas tests académiques canoniques et classiques, mais aussi sur des matrices publiques fournies par nos partenaires industriels privilégiés qui sont Airbus Group Innovations et le CEA-CESTA.

L'ensemble de ces travaux vise à répondre au défi majeur de concevoir et réaliser des solveurs robustes numériquement sur des supports d'exécution capables de passer à l'échelle et de repousser les limites des codes industriels existants en utilisant pleinement l'ensemble des ressources calculatoires telles que les CPUs, les GPUs et les Xeon Phi.

Les travaux correspondant à cette section sont regardés dans la thèse de G. Pichon et dans le travail postdoctoral de Y. Harness.

► Concurrence internationale.

Tous ces travaux seront menés dans un contexte de concurrence positive avec des équipes internationales renommées avec qui je souhaite développer des échanges réguliers, entre autres :

— SuperLU : Sherry Li, Lawrence Berkeley National Laboratory, USA

- UMFPACK (MATLAB) : Tim Davis, Florida University, USA
- TAUCS (Mathematica) : Sivan Toledo, Tel-Aviv University, IL
- WSMP (IBM) : Anshul Gupta, Minnesota University, USA
- PARDISO (INTEL) : Olaf Schenk, Basel University, CH
- HSL (NAG) : Jennifer Scott, Rutherford Appleton, UK
- BCSLIB (BOEING) : Daniel Pierce, Boeing, USA

Pour finir, je m'engagerai dans les réflexions au niveau d'Inria pour proposer un environnement unifié à l'ensemble des travaux autour des solveurs linéaires (creux) en partenariat avec les grands comptes du HPC (Airbus, EDF, CEA, TOTAL, IFP, ...).

Mon projet de recherche s'intègre clairement dans les nouvelles activités de l'équipe HiePACS et je me positionne comme un partenaire reconnu vis-à-vis des activités de plusieurs équipes Inria (ALPINE, ROMA, NACHOS) et de l'IRIT/CERFACS.

Liste des publications

Revues internationales :

- [1] G. Pichon, E. Darve, M. Faverge, P. Ramet, and J. Roman. *Sparse Supernodal Solver Using Block Low-Rank Compression : design, performance and analysis*. International Journal of Computational Science and Engineering, 27 :255-270, 2018.
- [2] G. Pichon, M. Faverge, P. Ramet, and J. Roman. *Reordering strategy for blocking optimization in sparse linear solvers*. SIAM Journal on Matrix Analysis and Application, 38(1) :226-248, 2017.
- [3] S. Moustafa, I. Dutka-Malen, L. Plagne, A. Poncot, and P. Ramet. *Shared Memory Parallelism for 3D Cartesian Discrete Ordinates Solver*. Annals of Nuclear Energy, 2014.
- [4] O. Coulaud, L. Giraud, P. Ramet, and X. Vasseur. *Developments in Parallel, Distributed, Grid and Cloud Computing for Engineering*. Chapter Augmentation and Deflation in Krylov subspace methods, pages 249-275. Saxe-Coburg Publications, Kippen, Stirlingshire, United Kingdom, 2013.
- [5] M. Barrault, B. Lathuilière, P. Ramet et J. Roman. *Efficient Parallel Resolution of The Simplified Transport Equations in Mixed-Dual Formulation*. Journal of Computational Physics, 230(5) :2004-2020, 2011.
- [6] G. Huysmans, Pamela S., E. van der Plas et P. Ramet. *Non-Linear MHD simulations of Edge Localised Modes (ELMs)*. Journal on Plasma Physics and Controlled Fusion, 51(12) :124012, 2009.
- [7] R. Abgrall, R. Huart et P. Ramet. *Numerical simulation of unsteady MHD flows and applications*. MagnetoHydroDynamics Journal, 45(2) :225-232, 2009.
- [8] P. Hénon, P. Ramet et J. Roman. *On finding approximate supernodes for an efficient ILU(k) factorization*. Parallel Computing, 34 :345-362, 2008.
- [9] P. Hénon, P. Ramet, et J. Roman. *PaStiX : A High-Performance Parallel Direct Solver for Sparse Symmetric Definite Systems*. Parallel Computing, 28(2) :301-321, 2002.
- [10] E. Caron, S. Chaumette, S. Contassot-Vivier, F. Desprez, E. Fleury, C. Gomez, M. Goursat, E. Jeannot, D. Lazure, F. Lombard, J.M. Nicod, L. Philippe, M. Quinson, P. Ramet, J. Roman, F. Rubi, S. Steer, F. Suter et G. Utard. *Scilab to Scilab//, the OURAGAN Project*. Parallel Computing, 27(11) :1497-1519, 2001.
- [11] D. Goudin, P. Hénon, F. Pellegrini, P. Ramet, J. Roman et J.-J. Pesqué. *Parallel Sparse Linear Algebra and Application to Structural Mechanics*. Numerical Algorithms volume 24, pages 371-391, 2000.

Congrès internationaux avec actes (Springer LNCS, IEEE ou SIAM) :

- [12] G. Pichon, E. Darve, M. Faverge, P. Ramet, and J. Roman. *Sparse Supernodal Solver Using Block Low-Rank Compression*. 18th IEEE International Workshop on Parallel and Distributed Scientific and Engineering Computing (PDSEC 2017), Orlando, USA, juin 2017.
- [13] S. Moustafa, M. Faverge, L. Plagne, P. Ramet. *3D Cartesian Transport Sweep for Massively Parallel Architectures with PARSEC*. 29th IEEE International Parallel & Distributed Processing Symposium (IPDPS'15), pages 581-590, Hyderabad, India, mai 2015.
- [14] A. Casadei, P. Ramet, and J. Roman. *An improved recursive graph bipartitioning algorithm for well balanced domain decomposition*. 21st IEEE International Conference on High Performance Computing, pages 1-10, Goa, India, décembre 2014.
- [15] X. Lacoste, M. Faverge, P. Ramet, S. Thibault, and G. Bosilca. *Taking advantage of hybrid systems for sparse direct solvers via task-based runtimes*. HCW'2014 workshop of IPDPS, pages 29-38, Phoenix, USA, mai 2014.

- [16] G. Huysmans, Pamela S., E. van der Plas et P. Ramet. *Non-Linear MHD simulations of Edge Localised Modes*. 36th EPS Plasma Physics Conference, Sofia, Bulgarie, juin 2009.
- [17] M. Barrault, B. Lathuilière, P. Ramet et J. Roman. *A domain decomposition method applied to the simplified transport equations*. IEEE 11th International Conference on Computational Science and Engineering, Sao Paulo, Brazil, pages 91-97, juillet 2008.
- [18] Y. Caniou, J.-S. Gay et P. Ramet. *Tunable parallel experiments in a GridRPC framework : application to linear solvers*. VECPAR'08, LNCS 5336, pages 430-436, Toulouse, France, juin 2008.
- [19] N. Kushida, Y. Suzuki, N. Teshima, N. Nakajima, Y. Caniou, M. Dayde et P. Ramet. *Toward an International Sparse Linear Algebra Expert System by Interconnecting the ITBL Computational Grid with the Grid-TLSE Platform*. VECPAR'08, LNCS 5336, pages 424-429, Toulouse, France, juin 2008.
- [20] M. Faverge et P. Ramet. *Dynamic Scheduling for sparse direct Solver on NUMA architectures*. Proceedings of PARA'08, Trondheim, Norway, à paraître dans LNCS, mai 2008.
- [21] P. Hénon, P. Ramet et J. Roman. *Partitioning and Blocking Issues for a Parallel Incomplete Factorization*. PARA'06, Workshop on state-of-the-art in scientific computing, Umea, Suède, LNCS 4699, pages 929-937, juin 2006.
- [22] P. Hénon, P. Ramet et J. Roman. *On using an hybrid MPI-Thread programming for the implementation of a parallel sparse direct solver on a network of SMP nodes*. Sixth International Conference on Parallel Processing and Applied Mathematics, Workshop HPC Linear Algebra, Poznan, Pologne, LNCS 3911, pages 1050-1057, september 2005.
- [23] P. Hénon, F. Pellegrini, P. Ramet, J. Roman, et Y. Saad. *Applying parallel direct solver skills to build robust and highly performant preconditioners*. PARA'04, Workshop on state-of-the-art in scientific computing, Copenhagen, Danemark, LNCS 3732, pages 601-619, juin 2004.
- [24] O. Beaumont, P. Ramet et J. Roman. *Asymptotically optimal algorithm for Laplace task graphs on heterogeneous platforms*. Fifth International Conference on Parallel Processing and Applied Mathematics (PPAM), Czesochowa, Pologne, LNCS 3019, pages 880-887, septembre 2003.
- [25] P. Hénon, P. Ramet et J. Roman. *Efficient algorithms for direct resolution of large sparse system on clusters of SMP nodes*. SIAM Conference LA'2003, Williamsburg, USA, juillet 2003.
- [26] P. Hénon, P. Ramet et J. Roman. *PaStiX : A Parallel Direct Solver for Sparse SPD Matrices based on Efficient Static Scheduling and Memory Managment*. SIAM Conference PPSC'2001, Portsmouth, Virginie, USA, mars 2001.
- [27] P. Hénon, P. Ramet et J. Roman. *PaStiX : A Parallel Sparse Direct Solver Based on a Static Scheduling for Mixed 1D/2D Block Distributions*. IPDPS'2000, Cancun, Mexique, LNCS 1800, pages 519-525, mai 2000.
- [28] P. Hénon, P. Ramet et J. Roman. *A Mapping and Scheduling Algorithm for Parallel Sparse Fan-In Numerical Factorization*. EuroPar'99, Toulouse, France, LNCS 1685, pages 1059-1067, septembre 1999.
- [29] F. Desprez, P. Ramet et J. Roman. *Optimal Grain Size Computation for Pipelined Algorithms*. EuroPar'96, Lyon, France, LNCS 1123, pages 165-172, septembre 1996.

Congrès internationaux avec comité de sélection :

- [30] M. Faverge, G. Pichon, P. Ramet, and J. Roman. *Exploiting Parameterized Task-graph in Sparse Direct Solvers* SIAM Conference on Computation Science and Engineering, Spokane, USA, février 2019.
- [31] G. Pichon, E. Darve, M. Faverge, P. Ramet, and J. Roman. *Block Low-rank Algebraic Clustering for Sparse Direct Solvers* SIAM Conference on Computation Science and Engineering, Spokane, USA, février 2019.
- [32] G. Pichon, E. Darve, M. Faverge, P. Ramet, and J. Roman. *Block Low-rank Algebraic Clustering for Sparse Direct Solvers* PMAA'2018, Zurich, Suisse, juin 2018.

- [33] G. Pichon, E. Darve, M. Faverge, P. Ramet, and J. Roman. SIAM Conference on Computation Science and Engineering, Atlanta, USA, février 2017.
- [34] G. Pichon, M. Faverge, and P. Ramet. *Exploiting Modern Manycore Architecture in Sparse Direct Solver with Runtime Systems*. SIAM Conference on Computation Science and Engineering, Atlanta, USA, février 2017.
- [35] G. Pichon, M. Faverge, P. Ramet, and J. Roman. *Impact of Blocking Strategies for Sparse Direct Solvers on Top of Generic Runtime*. SIAM Conference on Computation Science and Engineering, Atlanta, USA, février 2017.
- [36] E. Darve, M. Faverge, G. Pichon, P. Ramet, and J. Roman. *Sparse Supernodal Solver Using Hierarchical Compression*. Workshop on Fast Direct Solvers, Purdue, USA, novembre 2016.
- [37] P. Ramet. *On the use of low rank approximations for sparse direct solvers*. SIAM Annual Meeting, Boston, USA, juillet 2016.
- [38] M. Faverge, G. Pichon, P. Ramet, and J. Roman. *Impact of Blocking Strategies for Sparse Direct Solvers on Top of Generic Runtimes*. SIAM Conference on Parallel Processing for Scientific Computing, Paris, France, avril 2016.
- [39] E. Darve, M. Faverge, G. Pichon, P. Ramet, and J. Roman. *Exploiting H-Matrices in Sparse Direct Solvers*. SIAM Conference on Parallel Processing for Scientific Computing, Paris, France, avril 2016.
- [40] M. Faverge, G. Pichon, P. Ramet, and J. Roman. *Blocking strategy optimizations for sparse direct linear solver on heterogeneous architectures*. Sparse Days, Saint Girons, France, juin 2015.
- [41] M. Faverge, G. Pichon, P. Ramet, and J. Roman. *On the use of H-Matrix Arithmetic in PaStiX : a Preliminary Study*. Workshop on Fast Solvers, Toulouse, France, juin 2015.
- [42] X. Lacoste, M. Faverge, and P. Ramet. *A task-based sparse direct solver suited for large scale hierarchical/heterogeneous architectures*. SIAM Conference on Computation Science and Engineering, Salt Lake City, USA, février 2015.
- [43] A. Casadei, P. Ramet, and J. Roman. *Towards a recursive graph bipartitioning algorithm for well balanced domain decomposition*. SIAM Conference on Computation Science and Engineering, Salt Lake City, USA, février 2015.
- [44] P. Ramet. *On the design of parallel linear solvers for large scale problems*. International Congress on Industrial and Applied Mathematics, Pekin, China, aout 2015.
- [45] A. Casadei and P. Ramet. *Towards a recursive graph bipartitioning algorithm for well balanced domain decomposition*. International Congress on Industrial and Applied Mathematics, Pekin, China, aout 2015.
- [46] S. Moustafa, M. Faverge, L. Plagne, and P. Ramet. *Parallel 3D Sweep Kernel with PARSEC*. 16th IEEE International Conference on High Performance and Communications, workshop on HPC-CFD in Energy/Transport Domains, Paris, France, aout 2014.
- [47] A. Casadei, P. Ramet, and J. Roman. *Nested Dissection with Balanced Halo*. SIAM Workshop on Combinatorial Scientific Computing, Lyon, France, juillet 2014.
- [48] E. Agullo, M. Faverge, L. Giraud, A. Guermouche, P. Ramet, and R. Roman. *Toward parallel scalable linear solvers suited for large scale hierarchical parallel platforms*. WCCM-ECCM-ECFD, Barcelona, Spain, juillet 2014.
- [49] S. Moustafa, I. Dutka-Malen, L. Plagne, A. Poncot, and P. Ramet. *Shared Memory Parallelism for 3D Cartesian Discrete Ordinates Solver..* Joint International Conference on Supercomputing in Nuclear Applications + Monte Carlo, Paris, France, octobre 2013.
- [50] X. Lacoste, M. Faverge, and P. Ramet. *Sparse Linear Algebra over DAG Runtimes*. SIAM Conference on Computation Science and Engineering, Boston, USA, février 2013.
- [51] A. Casadei, L. Giraud, P. Ramet, and J. Roman. *Towards Domain Decomposition with Balanced Halo*. Workshop Celebrating 40 Years of Nested Dissection, Waterloo, Canada, juillet 2013.

- [52] P. Ramet. *From hybrid architectures to hybrid solvers*. Workshop Celebrating 40 Years of Nested Dissection, Waterloo, Canada, juillet 2013.
- [53] X. Lacoste, P. Ramet, M. Faverge, I. Yamazaki, G. Bosilca. *Toward a supernodal sparse direct solver over DAG runtimes*. PMAA'2012, London, England, juin 2012.
- [54] A. Casadei et P. Ramet. *Memory Optimization to Build a Schur Complement*. SIAM Conference LA'2012, Valencia, Spain, juin 2012.
- [55] X. Lacoste et P. Ramet. *Sparse direct solver on top of large-scale multicore systems with GPU accelerators*. SIAM Conference LA'2012, Valencia, Spain, juin 2012.
- [56] M. Faverge et P. Ramet. *Fine Grain Scheduling for Sparse Solver on Manycore Architectures*. SIAM Conference PPSC'2012, Savannah, USA, février 2012.
- [57] Y. Suzuki, N. Kushida, T. Tatekawa, N. Teshima, Y. Caniou, R. Guivarch, M. Dayde et P. Ramet. *Development of an International Matrix-Solver Prediction System on a French-Japanese International Grid Computing Environment*. Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo 2010, Tokyo, Japan, octobre 2010.
- [58] M. Barrault, B. Lathuilière, P. Ramet et J. Roman. *A Non Overlapping Parallel Domain Decomposition Method Applied to The Simplified Transport Equations*. International Conference on Mathematics, Computational Methods and Reactor Physics, New-York, USA, mai 2009.
- [59] R. Abgrall, O. Coulaud, P. Hénou, R. Huart, G. Huysmans, G. Latu, B. Nkonga, S. Pamela et P. Ramet. *Numerical simulation of tokamak plasmas*. 7th PAMIR International Conference on Fundamental and Applied MHD, Presqu'île de Giens, France, septembre 2008.
- [60] M. Faverge, X. Lacoste et P. Ramet. *A NUMA Aware Scheduler for a Parallel Sparse Direct Solver*. PMAA'2008, Neuchatel, Suisse, juin 2008.
- [61] M. Barrault, B. Lathuilière, P. Ramet et J. Roman. *A Domain Decomposition Method Applied to Large Eigenvalue Problems in Neutron Physics*. PMAA'2008, Neuchatel, Suisse, juin 2008.
- [62] M. Barrault, B. Lathuilière, P. Ramet et J. Roman. *A domain decomposition method for the resolution of an eigenvalue problem in neutron physics*. International Symposium on Iterative Methods in Scientific Computing (IMACS), Lille, France, mars 2008.
- [63] P. Hénou, P. Ramet et J. Roman. *On finding approximate supernodes for an efficient ILU(k) factorization*. PMAA'2006, Rennes, France, septembre 2006.
- [64] B. Braconnier, B. Nkonga, M. Papin, P. Ramet, M. Riccuto, J. Roman et R. Abgrall. *Efficient solution technique for low Mach number compressible multiphase problems*. PMAA'2006, Rennes, France, septembre 2006.
- [65] P. Hénou, F. Pellegrini, P. Ramet et J. Roman. *Blocking Issues for an Efficient Parallel Block ILU Preconditioner*. SIAM Conference On Preconditioning Techniques For Large Sparse Matrix Problems In Scientific And Industrial Applications, Atlanta, USA, mai 2005.
- [66] P. Hénou, P. Ramet et J. Roman. *A Blockwise Algorithm for Parallel Incomplete Cholesky Factorization*. PMAA'2004, Marseille, France, octobre 2004.
- [67] P. Hénou, B. Nkonga, P. Ramet et J. Roman. *Using of the High Performance Sparse Solver PaStiX for the Complex Multiscale 3D Simulations performed by the FluidBox Fluid Mechanics Software*. PMAA'2004, Marseille, France, octobre 2004.
- [68] P. Hénou, F. Pellegrini, P. Ramet, J. Roman et Y. Saad. *High Performance Complete and Incomplete Factorizations for Very Large Sparse Systems by using Scotch and PaStiX softwares*. SIAM Conference PPSC'2004, San Francisco, USA, février 2004.
- [69] P. Hénou, F. Pellegrini, P. Ramet et J. Roman. *Towards High Performance Hybrid Direct-Iterative Solvers for Large Sparse Systems*. SIAM Conference On Preconditioning Techniques For Large Sparse Matrix Problems In Scientific And Industrial Applications, Napa, USA, octobre 2003.

- [70] P. Hénon, P. Ramet et J. Roman. *Parallel factorization of very large sparse SPD systems on a network of SMP nodes*. PMAA'2002, Neuchâtel, Suisse, novembre 2002.
- [71] P. Hénon, P. Ramet et J. Roman. *PaStiX : A High-Performance Parallel Direct Solver for Sparse Symmetric Definite Systems*. PMAA'2000, Neuchâtel, Suisse, août 2000.
- [72] D. Goudin, P. Hénon, F. Pellegrini, P. Ramet, J. Roman et J.-J. Pesqué. *Description of the EMILIO Software Processing Chain and Application to Structural Mechanics*. PMAA'2000, Neuchâtel, Suisse, août 2000.

Ateliers internationaux sur invitation :

- [73] P. Ramet. *Heterogeneous architectures, Hybrid methods, Hierarchical matrices for Sparse Linear Solvers*. Seminar at Stanford, avril 2018.
- [74] P. Ramet. *From hybrid architectures to hybrid solvers*. Seminar at Stanford, juillet 2013.
- [75] P. Ramet. *Hybrid methods, Hybrid architectures, Hybrid compressions for sparse direct solvers*. Seminar at Stanford, novembre 2013.
- [76] P. Ramet. *Dynamic Scheduling for Sparse Direct Solver on NUMA and Multicore Architectures*. ComplexHPC meeting, Lisbon, Portugal, octobre 2009.
- [77] P. Hénon, P. Ramet et J. Roman. *A supernode amalgamation algorithm for an efficient block incomplete factorization*. Workshop on parallel iterative solvers and domain decomposition techniques, Minneapolis, USA, juillet 2008.
- [78] P. Hénon, P. Ramet et J. Roman. *A supernode amalgamation algorithm for an efficient block incomplete factorization*. Workshop PPAM'07, Gdansk, Pologne, septembre 2007.
- [79] P. Ramet. *High performances methods for solving large sparse linear systems - Direct and Incomplete Factorization*. Workshops ReDIMsOPS, Japan Atomic Energy Agency, Tokyo, Japon, mai 2007.
- [80] O. Czarny, G. Huysmans, P. Hénon et P. Ramet. *Improvement of existing solvers for the simulation of MHD instabilities*. Numerical flow models for controlled fusion, Porquerolles, France, avril 2007.
- [81] P. Hénon, F. Pellegrini, P. Ramet et J. Roman. *An efficient hybrid MPI/Thread implementation on a network of SMP nodes for the parallel sparse direct solver PaStiX : ordering / scheduling / memory management / out-of-core issues, and application to preconditioning*. Sparse Days and Grid Computing, Saint Girons, France, juin 2003.

Ateliers nationaux sur invitation :

- [82] M. Faverge, G. Pichon, P. Ramet, and J. Roman. *Utilisation de la compression low-rank pour réduire la complexité du solveur PaStiX*. JCAD'2018 : Journées Calcul et Données, Lyon, France, octobre 2018.
- [83] G. Pichon, E. Darve, M. Faverge, P. Ramet, and J. Roman. *Utilisation de la compression Block Low-Rank pour accélérer un solveur direct creux supernodal*. COMPAS'2018 - SOLHAR final meeting, Toulouse, France, juillet 2018.
- [84] P. Ramet. *Solveurs Directs*. Maison de la Simulation : Formation en Algèbre Linéaire Creuse Parallèle, Montpellier, France, mars 2015.
- [85] P. Ramet. *PaStiX : Parallel Sparse Matrix Package*. JDEV2015 : Journées Développement Logiciel, Bordeaux, France, juillet 2015.
- [86] P. Ramet. *On the design of parallel linear solvers for large scale problems*. Formation CNRS, Journée problème de Poisson, Paris, France, janvier 2015.
- [87] P. Ramet. *Task-based linear solvers for modern architectures*. 7th ITER International School, High Performance Computing in Fusion Science, Aix-en-Provence, France, aout 2014.

- [88] P. Ramet. *Solveurs Directs*. Maison de la Simulation : Formation en Algèbre Linéaire Creuse Parallèle, Paris, France, mars 2014.
- [89] P. Ramet. *Hybrid methods, Hybrid architectures, Hybrid compressions for sparse direct solvers*. Journée Scientifique du MCIA, février 2014.
- [90] P. Ramet. *Solveurs Directs*. Maison de la Simulation : Formation en Algèbre Linéaire Creuse Parallèle, Paris, France, mars 2013.
- [91] P. Ramet. *Méthodes directes et hybrides pour des solveurs creux adaptés aux machines multiC-PUs/multiGPUs*. 3ième Ecole Thématique de Simulation Numérique, Frejus, France, juillet 2013.
- [92] P. Ramet. *Sparse direct solver on top of large-scale multicore systems with GPU accelerators*. CEMRACS'2012, Méthodes numériques et algorithmes pour architectures pétaflopiques, Marseille, France, aout 2012.
- [93] P. Ramet. *Solveurs Directs*. Maison de la Simulation : Formation en Algèbre Linéaire Creuse Parallèle, Bordeaux, France, novembre 2011.
- [94] P. Ramet. *Linear algebra and sparse direct methods*. Séminaires de l'école MFN 2011 sur les méthodes et algorithmes pour le calcul haute performance, Roscoff, France, juin 2011.
- [95] P. Ramet. *Ordonnancement dynamique dans le solveur PaStiX pour des machines NUMA et multicœurs*. Formation CNRS, Solveurs de systèmes linéaires de grande taille : les avancées récentes, Lyon, France, novembre 2010.
- [96] P. Ramet. *Formation Parallélisme*. CEMRACS'2010, Modèles Numériques pour la Fusion, Marseille, France, août 2010.
- [97] P. Ramet. *Résolution de Systèmes Linéaires, Algorithmes et Parallélisme*. Formation CNRS, Informatique Scientifique pour le Calcul, Sète, France, octobre 2008.
- [98] P. Ramet et J. Roman. *Méthodes directes hautes performances de résolution en algèbre linéaire creuse*. Ecole CEA-EDF-Inria sur le calcul scientifique intensif, Rocquencourt, France, novembre 2006.

Congrès nationaux avec comité de sélection et avec actes :

- [99] P. Hénon, P. Ramet. *Optimisation de l'occupation mémoire pour un solveur parallèle creux direct hautes performances de type supernodal*. RenPar'2002, Hamamet, Tunisie, avril 2002.
- [100] P. Hénon, P. Ramet. *PaStiX : Un solveur parallèle direct pour des matrices creuses symétriques définies positives basé sur un ordonnancement statique performant et sur une gestion mémoire efficace*. RenPar'2001, Paris, France, avril 2001.
- [101] P. Ramet. *Calcul de la Suite Optimale de Taille de Paquets pour la Factorisation de Cholesky*. RenPar'9, Lausanne, Suisse, pages 111–114, mai 1997.
- [102] P. Ramet. *Calcul de la Taille Optimale des Paquets pour les Algorithmes Macro-Pipelines*. RenPar'8, Bordeaux, France, pages 21–24, juin 1996.

Rapports de fin de contrat et publications diverses :

- [103] M. Alaya, M. Faverge, X. Lacoste, A. Péré-Laperne, J. Péré-Laperne, P. Ramet, and T. Terraz. *Simul'Elec and PASTIX interface specifications*. Algo'Tech, 2015.
- [104] M. Faverge, X. Lacoste, P. Ramet, and T. Terraz. *Etude de la factorisation directe hétérogène et de la factorisation incomplète sur solveur PaStiX appliquées à des systèmes issus de problèmes du CEA/CESTA*. CEA/DAM/CESTA, 2015.
- [105] M. Boulet, G. Meurant, D. Goudin, J.-J. Pesqué, M. Chanaud, L. Giraud, P. Hénon, P. Ramet et J. Roman. *Résolution des systèmes linéaires sur calculateurs pétaflopiques*. Revue CHOCS vol 41, revue scientifique et technique de la Direction des Applications Militaires, janvier 2012.

- [106] G. Caramel et P. Ramet. *Optimisation des performances des outils de calcul de neutronique des coeurs*. E.D.F. / SINETICS, 2007.
- [107] P. Hénon, P. Ramet et J. Roman. *Evaluation des performances de la version SMP du solveur PaStiX de la chaîne logicielle EMILIO dans l'environnement du code ODYSSEE du CESTA*. CEA/DAM/CESTA, 2005.
- [108] P. Hénon, F. Pellegrini, P. Ramet et J. Roman. *Etude sur l'applicabilité de méthodes itératives nouvelles aux problèmes du CESTA*. CEA/DAM/CESTA, 2004.
- [109] P. Hénon, P. Ramet et J. Roman. *Amélioration et Extension du Solveur Direct Parallèle pour Grandes Matrices Creuses du CESTA*. CEA/DAM/CESTA, 2003.
- [110] D. Lecas et P. Ramet. *Parallélisation du code MIRO*. CEA/DAM/CESTA, 2001.
- [111] D. Goudin, P. Hénon, F. Pellegrini, P. Ramet et J. Roman. *Mise en oeuvre d'une Bibliothèque d'Outils pour la Résolution par Méthode Directe de Grands Systèmes Linéaires Creux Symétriques Définis Positifs sur Machine Parallèle*. CEA/DAM/CESTA, 2000.

Co-encadrement de thèses :

- [112] M. Faverge. *Ordonnancement hybride statique-dynamique en algèbre linéaire creuse pour de grands clusters de machines NUMA et multi-coeurs*. PhD thesis, LaBRI, Université Bordeaux, Talence, France, décembre 2009.
- [113] B. Lathuilière. *Méthode de décomposition de domaine pour les équations du transport simplifié en neutronique*. PhD thesis, LaBRI, Université Bordeaux, Talence, France, janvier 2010.
- [114] X. Lacoste. *Scheduling and memory optimizations for sparse direct solver on multi-core/multi-gpu cluster systems*. PhD thesis, LaBRI, Université Bordeaux, Talence, France, février 2015.
- [115] A. Casadei. *Optimizations of hybrid sparse linear solvers relying on Schur complement and domain decomposition approaches*. PhD thesis, LaBRI, Université Bordeaux, Talence, France, octobre 2015.
- [116] S. Moustafa. *Massively Parallel Cartesian Discrete Ordinates Method for Neutron Transport Simulation*. PhD thesis, LaBRI, Université Bordeaux, Talence, France, décembre 2015.
- [117] G. Pichon. *On the use of low-rank arithmetic to reduce the complexity of parallel sparse linear solvers based on direct factorization techniques*. PhD thesis, LaBRI, Université Bordeaux, Talence, France, novembre 2018.

Logiciels :

- [118] P. Ramet. PaStiX 5.*. Un solveur haute performance pour la résolution de grands systèmes linéaires basé sur une approche directe supernodale. Disponible à partir de l'adresse suivante : <http://pastix.gforge.inria.fr/>. Dépôt APP effectué par Inria (IDN.FR.001.230016.000.S.C.2008.000.31235).
- [119] P. Ramet. PaStiX 6.0. Disponible à partir de l'adresse suivante : <https://gitlab.inria.fr/solverstack/pastix>.