
**Dossier de Candidature
aux fonctions de
Professeur des Universités
Poste 451 - UFR de Mathématiques-Informatique**

Stéphane GENAUD

Date : 16 mars 2012

TABLE DES MATIÈRES

Table des matières

1	Curriculum Vitæ	4
1.1	État civil	4
1.2	Formation Universitaire	4
1.3	Expérience professionnelle	5
1.4	Responsabilités principales	5
2	Activités de recherches	6
2.1	Détails des thèmes de recherche	6
2.1.1	Spécification de programmes data-parallèles	6
2.1.2	Parallélisation pour la géophysique	6
2.1.3	Grilles de calcul	7
2.1.4	Simulation	8
2.2	Programme de recherches au sein du LSIIT	9
2.3	Liste de publications	14
2.4	Activités scientifiques	16
2.4.1	Niveau international	16
2.4.2	Niveau national	17
2.4.3	Niveau local	18
2.4.4	Jurys de thèse	18
2.5	Encadrements	18
2.5.1	Thèses	18
2.5.2	Stages de DEA/Master	19
2.5.3	Autres	19
2.6	Synthèse quantitative	20
3	Activités d'enseignement	21
3.1	Contexte	21
3.2	Enseignements en informatique	21
3.3	Tableau récapitulatif	22
3.4	Contenus des enseignements	23
3.5	Principales charges d'intérêt collectif	24
3.6	Projet pédagogique	24
3.6.1	Relations avec le monde professionnel	25
3.6.2	Internationalisation	25
4	Documents annexes	27
5	Copies d'articles	47

1 Curriculum Vitæ

1.1 État civil

Prénom et nom : **Stéphane GENAUD**
Date et lieu de naissance : Né le 10 mars 1969 à Forbach (57).
Nationalité : Française.
Situation familiale : Marié, 2 enfants.
Adresse professionnelle : LSIIT, Pôle API, Boulevard S. Brant,
F-67400 Illkirch
Tél. : +33 (0)38 92 77 449
Fax : +33 (0)3 90 24 45 47
Adresse électronique : genaud@unistra.fr
Page personnelle : <http://icps.u-strasbg.fr/members/genaud>

1.2 Formation Universitaire

2009 Habilitation à diriger des recherches en informatique

Université Henri Poincaré, Nancy. *Soutenue le 8/12/2009*

Titre du mémoire : *Exécutions de programmes parallèles à passage de messages sur grille de calcul.*

Jury :

- Pascal Bouvry (PU, Université du Luxembourg), Examineur,
- Christophe Cérin (PU, Université Paris 13), Rapporteur,
- Frédéric Desprez (DR, INRIA Grenoble Rhône-Alpes), Rapporteur,
- Claude Godart, (PU, Université Henri Poincaré), Président,
- Jens Gustedt (DR, INRIA Nancy Grand-Est), Garant,
- Thierry Priol (DR, INRIA Rennes Bretagne-Atlantique), Rapporteur.

1997 Doctorat en Sciences, mention Informatique

Université Louis Pasteur, Strasbourg.

Titre du mémoire : *Transformations de programmes PEI : applications au parallélisme de données.*

Jury :

- Luc Bougé (PU, ÉNS Lyon), Rapporteur,
- Christian Lengauer (PU, Université Passau, Allemagne), Examineur,
- Catherine Mongenet (PU, Université Louis Pasteur), Rapporteur interne,
- Guy-René Perrin (PU, Université Louis Pasteur), Directeur de thèse,
- Patrice Quinton (PU, Université de Rennes 1), Rapporteur.

1993 Diplôme d'Études Supérieures Spécialisées (DESS) en Informatique du parallélisme, Université de Franche-Comté, Besançon. Mention *très bien* (major).

1991 Bachelor of Sciences (BSc) in European Informatics, Sheffield Hallam University.

1.3 Expérience professionnelle

Seules les expériences liées à une activité d'enseignement et/ou de recherche sont décrites ici.

- 2007–2009 Détaché Chargé de Recherche** au Laboratoire Lorrain de Recherche en Informatique et ses Applications (LORIA, UMR 7503 CNRS-INPL-INRIA-UHP-Nancy 2), équipe projet INRIA ALGORILLE.
- 1998–2007 Maître de conférences en Informatique** à l'Université Robert Schuman, Strasbourg :
- ▷ Enseignant à l'IECS, Université Robert Schuman,
 - ▷ Chercheur au LSIIT, UMR 7005 CNRS-ULP, équipe ICPS.
- 1996–1998 Attaché Temporaire d'Enseignement et de Recherche** au département d'informatique de l'IUT de l'université Robert Schuman.
- 1994–1996** Doctorant à l'Université Louis Pasteur (ULP), Strasbourg.
Directeur de thèse : Guy-René Perrin.
- 1993–1994** Doctorant à l'Université de Franche-Comté, Besançon.
Directeur de thèse : Guy-René Perrin.

Position administrative actuelle (depuis septembre 2009)

- **Maître de conférences en Informatique** à l'Université de Strasbourg (UdS) :
 - Enseignant à l'École de Management Strasbourg,
 - Chercheur au Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection (LSIIT, UMR 7005 CNRS-UdS). équipe *Image et Calcul Parallèle Scientifique* (ICPS).
- Titulaire de la **Prime d'Excellence Scientifique** depuis septembre 2009.
- **Qualifié aux fonctions de Professeur des Universités** dans la section 27 en 2010.
Numéro de qualification : PR-2010-27-10127207589

1.4 Responsabilités principales

- **Responsabilités pédagogiques**
 - Directeur délégué aux Systèmes d'Information de l'École de Management Strasbourg depuis 07/2011.
 - Responsable de la filière *Système d'Information* du master Management International de l'IECS de 2002 à 2007.
- **Responsabilités recherche**
 - Responsable du thème de recherche *Grilles* dans l'équipe ICPS du LSIIT de 2001 à 2007. Reprise du thème à mon retour dans l'équipe en septembre 2009.
 - Membre du conseil scientifique du département *Expertise pour la recherche de l'UdS* depuis 2010.
- **Responsabilités administratives et collectives**
 - Membre de 3 commissions de spécialistes entre 2004 et 2008.
 - Membre du comité d'experts UdS (section 27) depuis 2009 et membre dans 2 comités de sélection externes.
 - Co-responsable des systèmes d'information de mon établissement depuis 2001.

2 Activités de recherches

Résumé *Mes thèmes de recherche concernent le parallélisme, essentiellement sur des architectures de type cluster, grilles ou cloud. Ma thèse de doctorat avait pour objet l'écriture et la transformation de programmes parallèles à l'aide d'un langage formel. A la suite de cette approche formelle, je me suis investi sur un problème réel dans le domaine de la géophysique, nécessitant la conception et le développement d'applications parallèles. Lorsqu'ont émergé les grilles, j'ai étudié certains problèmes nouveaux que posaient ces systèmes hétérogènes, et la façon dont des applications pouvaient y être déployées. Mon travail sur les grilles a concerné l'évaluation et l'amélioration des performances des programmes parallèles dans ce contexte, puis l'amélioration des logiciels pour mieux prendre en charge les programmes parallèles. L'utilisation des principes des systèmes pair-à-pair pour la découverte et l'auto-organisation des ressources ainsi que des mécanismes de tolérance aux pannes par réplication des calculs ont été proposés. Parallèlement aux évaluations sur des cas et des plate-formes réelles, je travaille à la simulation de programmes distribués en contribuant à l'outil SIMGRID. Je m'intéresse particulièrement à l'extension du simulateur pour permettre la simulation de programmes MPI sans modification du code source. Enfin, mes recherches actuelles sont tournées vers les problèmes d'allocation des ressources et d'ordonnancement des tâches, dans le but de proposer aux clients de bons compromis performance/prix sur des plates-formes virtualisées comme celles fournies par les clouds IaaS.*

Ces différents aspects sont détaillés ci-après, puis suivent la liste de mes publications, mon activité d'animation scientifique, mes participations à des encadrements, et mes perspectives de recherches.

2.1 Détails des thèmes de recherche

2.1.1 Spécification de programmes data-parallèles

Ma thèse de doctorat [2] a porté sur la définition et l'utilisation d'un langage formel baptisé Pei [6]. Ce formalisme permet la description de programmes pour des ordinateurs parallèles, dans un modèle de programmation de type *parallélisme de données*. Nous avons montré comment ce formalisme pouvait être utilisé pour raisonner sur les programmes et les transformer en nouveaux programmes sémantiquement équivalents ou raffinés [29, 26, 25, 10]. Le deuxième volet du travail a consisté à proposer des méthodes pour traduire ces énoncés formels vers des langages parallèles cibles comme *High Performance Fortran* ainsi que les logiciels permettant les transformations, le contrôle de validité des programmes ainsi que les compilateurs pour la traduction¹.

2.1.2 Parallélisation pour la géophysique

Après cette expérience d'approche formelle d'un modèle de programmation pour le parallélisme, je me suis plongé dans un travail concret de parallélisation sur des codes scientifiques. Je me suis investi en particulier dans la conception et le développement d'outils logiciels pour la géophysique. L'enjeu est de pouvoir calculer une tomographie sismique globale en ondes de volumes permettant d'améliorer un modèle de vitesses des ondes sismiques, et de là en déduire des propriétés géologiques à l'intérieur de la Terre. L'ensemble de ces outils² conçus lors d'une thèse en collaboration avec l'Institut de Physique du Globe de Strasbourg (UMR CNRS-UdS 7516) permet de combiner et d'enchaîner des traitements sur

1. http://icps.u-strasbg.fr/pei/PEI_SUMMARY/langage.htm

2. <http://renass.u-strasbg.fr/ray2mesh>

[6] Eric Violard and Guy-René Perrin. PEI : a language and its refinement calculus for parallel programming. *Parallel Computing*, (18) :1167–1184, 1992.

des données géophysiques extrêmement volumineuses. Le travail s'est concrétisé par une tomographie utilisant la totalité des données (sismogrammes) acquises depuis 1965 par les réseaux de surveillances sismiques à travers le monde. Pour répondre à cet objectif les applications ont été conçues pour s'exécuter sur des architectures parallèles, et ont été testées sur des configurations très différentes allant de la machine parallèle à des réseaux de station de travail en passant par des grilles de calcul (cf. paragraphe suivant) à l'échelle nationale [9, 19].

2.1.3 Grilles de calcul

L'exploitation de telles applications scientifiques pose des problèmes concrets quant au choix de la machine cible pour l'exécution. Un type d'architecture nouveau a émergé ces dernières années grâce aux technologies qui permettent de fédérer efficacement des ressources de calcul provenant d'institutions différentes. Cette alternative a été popularisée ^[3,4] sous l'appellation de *grilles*. J'ai porté en 2001 un projet intitulé *Transformations et Adaptations de programmes pour la Grille* (TAG) accepté dans le cadre de l'Action Concertée Incitative (ACI) Grid du Ministère de la Recherche (2002–2005). Le projet visait l'étude du comportement d'applications scientifiques, essentiellement des programmes MPI ³ sur les grilles. Deux catégories de problèmes ont été traitées dans ce projet. D'une part, déterminer quelles *performances* on peut espérer d'une application exécutée sur des ressources hétérogènes distribuées à large échelle géographique et ce qu'on peut faire pour les améliorer. D'autre part, travailler sur la couche *intergicielle* pour masquer la complexité de ces systèmes répartis et améliorer la prise en charge des applications.

Performances Pour améliorer les performances, une partie de mon travail a été l'étude de techniques d'équilibrage de charge statiques ou dynamiques. En effet, les découpages des données et la distribution de la charge dans de nombreuses applications parallèles font l'hypothèse que les processeurs et les réseaux sont homogènes. Pour optimiser le temps de calcul d'un ensemble de tâches indépendantes exécutées sur des processeurs hétérogène, il faut déterminer la meilleure quantité de données distribuées à chacune. Nous avons proposé des algorithmes pour déterminer statiquement de telles distributions optimales. Ces algorithmes ont été testés sur de vraies applications, dont l'application de géophysique décrite précédemment [23, 22, 7]. Nous avons également comparé cette approche statique à l'approche dynamique dans laquelle le maître distribue un nouveau bloc de travail à un esclave dès que celui-ci le demande : l'équilibrage de charge se fait alors naturellement selon le rythme de calcul des esclaves. La difficulté dans cette approche est de déterminer la taille du bloc à envoyer : en dessous de la taille idéale, il y a des aller-retours inutiles, tandis qu'au dessus, les esclaves ne finissent pas simultanément. En revanche, cette stratégie possède l'avantage majeur de s'adapter aux variations de charge.

Intergiciel Les travaux précédents se sont largement appuyés sur l'expérimentation réelle sur des grilles construites avec le logiciel Globus. La prise en charge de nos applications par cet intergiciel ne correspondant pas à nos attentes, nous avons proposé un nouveau type d'intergiciel spécialisé pour MPI. Parmi les lacunes observées pour l'exécution des programmes parallèles figurent l'absence de mécanisme de co-allocation de ressources sur différents sites, la détection de la disponibilité des ressources

3. Message Passing Interface définit une bibliothèque de fonctions permettant à des processus d'échanger des messages. C'est devenu le standard *de facto* pour les programmes parallèles exécutés sur des architectures à mémoire distribuée.

[3] Ian Foster and Carl Kesselman. Globus : A metacomputing infrastructure toolkit. *The International Journal of Supercomputer Applications and High Performance Computing*, 11(2) :115–128, 1997.

[4] Ian Foster and Carl Kesselman, editors. *The Grid : Blueprint for a New Computing Infrastructure*. Morgan Kaufmann Publishers, August 1998.

au moment précis de l'exécution, la gestion de binaires multiples pour chaque systèmes, la difficulté d'accéder aux fichiers de données et programmes, et l'absence de détection des pannes et de tolérance aux pannes. Nous avons développé P2P-MPI⁴, pour pallier ces lacunes. P2P-MPI comprend à la fois la couche intergicielle et la bibliothèque de communication permettant de développer des programmes parallèles à passage de messages. Cette dernière est une implémentation de MPJ⁵ qui intègre la notion de tolérance aux pannes. Les applications sont des programmes Java, beaucoup plus faciles à déployer dans un environnement hétérogène. Pour la couche intergicielle, nous reprenons les principes des systèmes pair-à-pair : chaque machine démarrée avec P2P-MPI devient un pair susceptible de partager son CPU ou d'utiliser ceux des autres. Cette approche confère autonomie et robustesse aux applications. L'autonomie provient de la possibilité de découvrir dynamiquement un ensemble de pairs disponibles à ce moment précis pour exécuter un programme parallèle. On construit ainsi dynamiquement une "plate-forme" à chaque demande d'exécution. La façon de choisir les pairs les plus adaptés parmi ceux disponibles dépend de plusieurs critères, dont la latence réseau qui sépare la machine qui fait la requête des pairs candidats, la présence ou non des données nécessaires dans les caches des pairs distants, et le souhait de l'utilisateur de concentrer ou non les processus sur le minimum de machines.

Dans de tels systèmes, les pannes sont fréquentes. Or, lors d'un calcul, la panne d'un des participants provoque l'arrêt de l'application. Pour diminuer le risque de panne, nous avons proposé une solution jamais expérimentée dans ce contexte qui est la réplication des calculs. L'utilisateur décide du taux de redondance de l'exécution de chaque processus, sur des machines différentes, et le système gère de manière transparente la cohérence de l'exécution. En cas de panne de l'un des processus, l'application peut poursuivre son exécution tant qu'il subsiste au moins une copie de ce processus de calcul.

La conception de P2P-MPI a été décrite dans [6, 20]. Nous avons aussi démontré sa capacité à prendre en charge l'exécution de programmes parallèles sur plusieurs centaines de processeurs [16]. La thèse de Choopan Rattanapoka [5] présente l'ensemble des résultats. P2P-MPI est aussi un support pour l'étude de la tolérance aux pannes. Nous avons étudié le mécanisme de réplication des calculs que nous proposons [18] en montrant comment déterminer un taux optimal de réplication [4] puis en faisant une étude quantitative du coût de la réplication et des temps de reprise [15].

Enfin, je me suis attaché à valider notre proposition sur de vraies applications. Deux collaborations ont abouties dans ce cadre. En 2007 et 2008 nous avons aidé des collègues du LSIIT (Pierre Gançarski), dans le domaine de la fouille de données, à paralléliser leur méthode d'apprentissage non-supervisée pour le clustering [5]. En 2008 et 2009, nous avons collaboré avec des collègues de SUPELEC (Virginie Galtier et Stéphane Vialle) pour comparer les implantations parallèles de la méthode d'apprentissage Adaboost dans deux modèles de programmation différents (JavaSpace et MPJ) [14].

2.1.4 Simulation

Les expérimentations que nous avons menées dans notre travail sur les grilles ont été fastidieuses. L'apparition de l'outil Grid'5000 a considérablement élargi les possibilités d'expérimentation dans un environnement réel, tout en permettant un protocole expérimental plus rigoureux car l'utilisateur peut sélectionner le matériel voulu, et installer le système d'exploitation de son choix. Le caractère reproductible des expériences a donc été considérablement amélioré avec Grid'5000, mais reste néanmoins imparfait, car le réseau reliant les sites ainsi que les clusters sont régulièrement renouvelés. Les expériences doivent donc se succéder sur une période relativement courte pour obtenir une série de résultats

4. <http://www.p2pmi.org>

5. Message Passing for Java : adaptation de MPI pour Java

[5] Choopan Rattanapoka. *P2P-MPI : A fault-tolerant Message Passing Interface Implementation for Grids*. PhD thesis, University Louis Pasteur, Strasbourg, April 2008.

comparables. Cependant, les problèmes techniques ou la difficulté d’obtenir l’accès à l’outil au moment souhaité peuvent fortement allonger la période d’expérimentation. La simulation présente face à ce problème un grand intérêt. Elle peut permettre de tester de nombreux scénarios, et de ne faire des expériences réelles que pour les cas les plus intéressants.

SIMGRID [2] est un projet important dans le paysage de la recherche académique sur la simulation des systèmes distribués. Le logiciel permet de décrire l’ensemble des ordinateurs connectés et le réseau les interconnectant, ainsi que les opérations de calcul et de communication survenant au sein d’une application, et d’en faire une simulation à événements discrets. Le déroulement d’une application est décrit à travers une *interface* au simulateur, c’est-à-dire une API fournie pour décrire ces opérations de calcul et de communication.

Ce logiciel a été soutenu entre autres par le projets ANR USS-SimGrid^{6 7} (2009-2011), puis aujourd’hui par le projet SONGS⁸ (2012-2015). Je participe à ces projets, dont l’objectif est d’étendre les capacités de l’outil. Il n’existait pas par exemple, avant le projet Uss-Simgrid, d’interface permettant de simuler les programmes parallèles à passage de messages (ses deux principales API proposaient alors des communications point à point bloquantes). Nous avons proposé une interface pour simuler des programmes MPI. J’ai réactivé un effort fait dans ce sens à l’université d’Hawaï (Mark Stillwell et Henri Casanova) mais inachevé, baptisée SMPI. Ce travail, poursuivi par Pierre-Nicolas Clauss a abouti à une version désormais livrée avec SimGrid. La conception et l’évaluation de SMPI ont fait l’objet d’une publication à la conférence IPDPS [13]. Plusieurs perspectives sont ouvertes avec cet outil. On peut extrapoler l’exécution d’un programme sur une machine qui n’existe pas encore, dont on ne donne que la description. Ceci peut servir par exemple à des fins de dimensionnement d’un cluster. Cela peut être utile dans de nombreuses autres situations, comme l’enseignement, où une machine de bureau peut servir à simuler un cluster ou un système distribué. On peut également imaginer dans le futur continuer des exécutions en simulation après avoir capturé la trace d’un programme réellement exécuté (avec des outils de profiling existants) et en injectant cette trace dans le simulateur.

2.2 Programme de recherches au sein du LSIIT

Mon projet pour la recherche s’inscrit naturellement dans la continuité et l’évolution de mon activité actuelle au Laboratoire des Sciences de l’Image, de l’Informatique et de la Télé-détection (LSIIT), dans l’équipe Image et Calcul Parallèle Scientifique (ICPS). Je souhaite développer mes activités de recherche dans ce laboratoire, et à partir de 2013 dans le laboratoire Icube qui succèdera à l’actuel LSIIT.

Ma stratégie de développement comporte deux axes : un renforcement de l’activité du thème *applications* de l’équipe, et une extension du thème *grilles* aux environnements de calcul basés sur l’externalisation des ressources, c’est-à-dire les *clouds*. Ces deux axes sont complémentaires.

Thème Applications

Les activités de l’équipe ICPS sont structurées en trois thèmes :

6. <http://uss-simgrid.gforge.inria.fr/>

7. Labelisé projet *phare* par l’ANR.

8. <http://infra-songs.gforge.inria.fr/>

[2] Henri Casanova, Arnaud Legrand, and Martin Quinson. SimGrid : a Generic Framework for Large-Scale Distributed Experiments. In *10th IEEE International Conference on Computer Modeling and Simulation*. IEEE Computer Society Press, March 2008.

1. la *compilation et l'optimisation de programmes*, qui cherche à exhiber du parallélisme qu'on peut mettre en œuvre à l'intérieur d'une machine (par exemple sur des multi-cœurs),
2. l'*adaptation des programmes pour les grilles*, qui étudie la mise en œuvre d'applications quand le parallélisme mobilise plusieurs machines interconnectées, potentiellement hétérogènes (par exemple une fédération de clusters à l'échelle nationale ou internationale),
3. les *applications*, qui à travers l'étude de cas applicatifs concrets dans le domaine scientifique, examine l'apport du parallélisme pour améliorer les résultats scientifiques envisageables par simulation numérique.

Ce dernier thème permet à l'ICPS d'établir des collaborations multi-disciplinaires et apporte des problématiques réelles aux deux autres thèmes. Ces problèmes sont variés dans leur forme mais sont tous relatifs à la simulation numérique. La simulation numérique est un besoin crucial dans de nombreuses disciplines scientifiques. L'augmentation théorique des capacités de calcul des ordinateurs ne répond que très partiellement aux exigences croissantes des utilisateurs. Ces derniers peuvent par exemple vouloir ajouter une dimension au problème, accroître la résolution, ou simuler sur une période plus longue, etc. Répondre à ces exigences demande d'exploiter au mieux les dispositifs matériels à disposition dans le domaine du calcul intensif. Cette tâche est difficile étant donné la diversité et la complexité du matériel et des couches système. On est aujourd'hui réduit à examiner une application comme un cas particulier, pour lequel on ne sait souvent pas prédire quelles configurations matérielles et logicielles vont permettre les meilleures performances : on doit évaluer l'impact de différents modèles de programmation, du nombre de cœurs par CPU, du réseau d'interconnexion entre CPU et mémoire, évaluer l'opportunité d'utiliser des GPU, faire de la vectorisation (SSE), etc.

Les exigences des applications de calcul intensif sont un moteur indéniable pour la recherche en parallélisme. Les problèmes posés déclenchent une réflexion sur les modèles de programmation, sur les outils automatiques d'analyse de code et les méthodes de compilation, sur les intergiciels pour coordonner l'exécution distribuée des processus ou la gestion des pannes, sur l'architecture, sur les méthodes numériques. Les deux derniers domaines ne sont pas à proprement parler dans notre champ de compétence mais doivent être maîtrisés, par exemple à travers des collaborations avec des spécialistes. Le travail que j'ai mené pour des applications en géophysique m'a ainsi beaucoup appris. Le code développé a souvent servi d'exemple dans la communauté travaillant sur Grid'5000 pour évaluer des solutions proposées pour des environnements de type grilles ^[1].

Aujourd'hui, je compte poursuivre ce travail interdisciplinaire, à travers une collaboration avec Guillaume Latu, ancien membre de l'équipe ICPS, maintenant employé au CEA Cadarache. Nous co-encadrons depuis février 2011 un doctorant, Matthieu Kuhn, avec Nicolas Crouseille chargé de recherches en mathématiques appliquées dans l'équipe CalVi⁹, pour paralléliser un code de physique de plasmas de bords. Ce travail renforce une collaboration de longue date (depuis 2002) entre l'ICPS et CalVi. Cette collaboration a été à l'origine de notre association à la proposition de LabEx avec l'IRMA. Ce LabEx a été obtenu en février 2012, et va donc permettre de renforcer de manière significative ces travaux interdisciplinaires, notamment à travers des thèses co-encadrées. Nous souhaitons également recruter un permanent sur ce thème. Un poste de CR2 CNRS ayant été fléché pour recruter un informaticien dans un laboratoire de mathématiques, nous avons travaillé en concertation avec l'IRMA pour porter

9. L'équipe-projet CalVi est bi-localisée entre Strasbourg et Nancy. Nos collaborateurs directs font partie du laboratoire de mathématiques IRMA (UMR 7501 CNRS-UdS).

[1] Franck Cappello and Henri E. Bal. Toward an international "computer science grid". In *Seventh IEEE International Symposium on Cluster Computing and the Grid (CCGrid 2007)*, pages 3–12, 2007.

un dossier de candidature calibré pour ces besoins à ce concours.

Je compte également démarrer de nouvelles collaborations, comme celle avec l’institut pluridisciplinaire Hubert Curien (UMR 7178, CNRS-IN2P3-INC-INEE, UdS) dans le domaine de la protéomique (Alain Van Dorsselaer). Après une première analyse des besoins en protéomique, nous pensons devoir concentrer la plupart de nos efforts sur le couplage entre calculs et données, avec le plus souvent des calculs très nombreux mais indépendants, et utilisant de grandes bases de données dont le contenu est mis à jour presque quotidiennement. Les environnements de grille sont appropriés à ce type de travail, étant donné les capacités de calcul qu’offrent l’organisation virtuelle BioMed (sous-ensemble des ressources de la grille européenne EGI). Cette collaboration est importante pour les protéomistes Strasbourgeois qui ont obtenu la responsabilité de co-développer (avec deux autres centres) la plate-forme ProFI¹⁰, sélectionnée dans l’appel à projets “Infrastructures nationales de recherche en biologie”. Trois étudiants sont en stage depuis février 2012 pour apporter des améliorations à la plate-forme développée à l’IPHC [30] ou faire un état de l’art des technologies de cloud et tester leur introduction dans cet environnement de production.

Thème Grilles et Clouds

A mon retour au LSIIT en septembre 2009, j’ai infléchi l’orientation de mes recherches vers les problématiques de *cloud* qui me semblent incontournables dans le domaine. En comparaison des grilles, par nature hétérogènes, les clouds proposent des ressources de calcul provenant de clusters hébergés dans des *data centers*, dont on peut décider du type (CPU, quantité mémoire). De plus, la virtualisation (cloud IaaS) permet de rendre également la configuration logicielle homogène, en transportant tout l’environnement logiciel dans la machine virtuelle exécutée chez le fournisseur de ressources. C’est donc un nouveau type d’architecture, entre le cluster homogène et la grille hétérogène, qui s’offre à nous. La disparition de l’hétérogénéité est un élément capital pour le succès de ce type de solution. On peut ainsi simplifier considérablement les intergiciels comparativement à ceux utilisés traditionnellement dans les grilles. Cependant, les problématiques fondamentales des grilles, à savoir l’allocation des ressources, l’ordonnancement des calculs et des données, la tolérance aux pannes, demeurent mais doivent être revues dans ce contexte. D’un point de vue économique, les fournisseurs de cloud proposent généralement une facturation proportionnelle au temps d’utilisation uniquement, ce qui rend cette solution très attractive car supprimant totalement le coût de possession des matériels. Les économies d’échelle réalisées grâce à la virtualisation et à la taille des *data centers* permettent par ailleurs aux fournisseurs de proposer des prix très compétitifs. On assiste aujourd’hui à une multitude de projets visant à évaluer la viabilité de cette approche, en particulier dans le calcul scientifique. Par exemple, dès 2009, le département de l’énergie américain a (DOE) décidé de financer le projet Magellan¹¹, infrastructure de 5000 coeurs exploitée sous la forme d’un cloud ouvert à 3000 scientifiques pour tester un large éventail d’applications scientifiques. Je crois qu’une expérimentation comparable serait profitable pour le méso-centre de calcul de l’université de Strasbourg, amené à jouer un rôle important après le succès de sa proposition d’Equipex. Nos relations privilégiées avec ce centre (Romaric David) encouragent un tel projet.

Faire émerger un thème de recherche nécessite cependant un investissement sur plusieurs années. Avec Julien Gossa, maître de conférences, nous avons initié de nombreuses démarches pour nous établir dans notre environnement local et national, afin d’être en position de travailler à des projets de visibilité internationale.

10. <http://bit.ly/eXXCOR>

11. <http://www.nersc.gov/nusers/systems/magellan/>

Au niveau local, la présence de ce thème dans nos enseignements a permis de recruter un étudiant de master en stage de recherche, puis en thèse en octobre 2011 sur un financement DGA.

Au niveau national, nous avons contribué à la proposition de projet ANR baptisée SONGS, continuant le projet USS-SimGrid (voir section 2.1.4). Notre proposition a été acceptée et SONGS a démarré début 2012. Ce projet qui implique 19 chercheurs sur 7 sites est de type *plateforme* au sens ANR, c'est-à-dire qu'il implique déjà une large communauté de chercheurs et d'utilisateurs très actifs, il est ouvert et son fonctionnement est durable. Le projet est aidé à hauteur de 1,8 M€. L'objectif est d'étendre et d'améliorer les capacités de l'outil de simulation SimGrid en prenant en compte l'évolution des architectures distribuées. Parmi les domaines visés, on compte les plates-formes de *volunteer computing*, les grilles de production, le HPC, et les clouds IaaS. Dans ce projet, je suis pour le LSIIT le responsable scientifique (porteur local) d'un work package sur la simulation des clouds.

Complémentarité des thèmes

La diversité des besoins en terme de calcul, face à la diversité croissante des architectures de calcul parallèle ou distribué, m'incite à mener de front des états de l'art sur ces deux thèmes. Je crois que l'équipe peut contribuer sur deux aspects : la performance des applications et la gestion des ressources.

Nous avons parlé de la performance dans le thème applicatif, qui nécessite une coopération étroite entre les chercheurs du domaine et les chercheurs expérimentés en parallélisme, ces derniers pouvant apporter des méthodes applicables directement au problème. Parmi ces méthodes, du ressort de la recherche en informatique, on compte l'analyse (par exemple de trace, de complexité, la visualisation), la parallélisation et l'optimisation automatique, l'expérimentation et la simulation.

La gestion des ressources nécessite de construire des environnements logiciels de type intergiciel, capables de recenser, sélectionner, configurer des ressources adaptées pour une application à un moment donné. Nous avons évoqué l'évolution des équipements de calcul vers plus de complexité. Les machines individuelles voient leur puissance de calcul continuer de croître grâce à l'augmentation du nombre de cœurs, tandis que l'accès à des équipements de calcul de type cluster est facilité car ils sont de plus en plus nombreux [7] avec une meilleure connectivité réseau. Certaines applications pourront être avantageusement prises en charge par une seule machine multi-cœurs, tandis que d'autres tireront plus de profit d'un petit laps de temps, éventuellement en mode *best-effort*, accordé sur un cluster avec beaucoup de processeurs. Enfin, la nouvelle voie ouverte par les clouds IaaS permet d'envisager d'embarquer l'application dans une image de machine virtuelle pour l'exécuter sur des ressources de calcul externalisées.

Choisir la bonne plate-forme d'exécution est donc un problème à part entière. Cette problématique est au centre de mon projet de recherches au LSIIT. De manière générale, je souhaite étudier les capacités des différents types de plates-formes, en intégrant, quand c'est possible, la dimension économique.

Optimiser l'exploitation des ressources revêt plusieurs formes. J'ai commencé à travailler dans le cadre du projet ANR SPADES (2009-2011)¹² sur un système permettant d'exploiter des tranches de temps réduites qui peuvent apparaître de manière fugace sur des clusters. L'exploitation de ces tranches se faisant en mode non-prioritaire (ou *best-effort*) on pourrait ainsi augmenter le taux d'utilisation de ces clusters. Pour atteindre cet objectif, il faut concevoir des intergiciels extrêmement réactifs,

12. <http://graal.ens-lyon.fr/SPADES/SPADES>

[7] Jie Wu, Earl C. Joseph, Richard Walsh, and Steve Conway. Worldwide technical computing server 2009–2013 forecast. Technical Report Doc #217232, IDC, 2009. <http://www.idc.com/getdoc.jsp?containerId=217232>.

capables de localiser rapidement sur un ensemble de clusters, les plages d'utilisation disponibles, ainsi que la compatibilité du système avec les besoins de l'application. Un deuxième travail démarré dans ce sens est l'étude de stratégies d'allocation puis d'ordonnancement *online* de jobs séquentiels sur une ou plusieurs plates-formes de type cloud IaaS. L'utilisation des ressources dans ces centres étant facturée par tranche de temps (typiquement toute heure entamée est due), de nombreux jobs laissent des périodes de temps facturées mais non utilisées, qui peuvent donc être utilisées gratuitement pour d'autres jobs. Une stratégie d'allocation doit décider pour chaque job si une nouvelle machine virtuelle doit être démarrée pour ce traitement ou si l'on peut réutiliser une machine virtuelle déjà démarrée. On étudie dans ce contexte un problème bi-critère de performance incluant le temps d'attente des jobs et le coût de l'ensemble des jobs. Ces résultats ont été présentés à la dernière édition de la conférence CLOUD 2011 [12]. L'objectif à plus long terme est de construire un système de courtage permettant de sélectionner, éventuellement auprès de différents fournisseurs, des machines virtuelles dont les capacités sont les mieux adaptées au besoin.

Conclusion

L'équipe ICPS souhaite maintenir l'équilibre dans le développement de ses différents thèmes. Je suis convaincu qu'il faut couvrir tout le spectre du parallélisme pour avoir une bonne vision des solutions les plus appropriées aux différents types de besoins. Cette stratégie me semble particulièrement importante dans un contexte où les technologies changent très vite. Certaines applications qui nécessitaient un petit cluster il y a cinq ans peuvent utiliser une machine multi-cœurs aujourd'hui. Ce changement d'architecture cible peut induire un changement quant au modèle de programmation ainsi que des outils d'optimisation utilisables. C'est pourquoi il me semble important d'avoir au sein de l'équipe l'ensemble des visions et des compétences du parallélisme. J'ai par exemple entamé un travail [31] avec Alain Ketterlin, sur l'abstraction des schémas de communications des programmes MPI. Ce travail basé sur la reconnaissance de répétitions linéaires s'appuie sur des résultats obtenus initialement pour l'analyse des accès mémoire, menés dans le thème compilation et optimisation de programmes.

En conclusion, je pense qu'un poste de professeur dans l'équipe ICPS amènerait une accélération du développement du thème grilles et clouds, ce statut ayant une plus grande reconnaissance au plan national ou international. Je souhaite également contribuer au développement du thème applications, avec l'espoir de pouvoir recruter un spécialiste à l'interface dans le moyen terme. Je pense pouvoir faire bénéficier le laboratoire et l'équipe, de ma connaissance de l'environnement local, importante pour les collaborations en cours ou à venir, et de mon intégration dans la communauté de recherche des grilles ou du HPC, avec en particulier le travail fait dans le projet SONGS amené à avoir une grosse visibilité internationale.

2.3 Liste de publications

Thèses

- [1] **Stéphane Genaud**. *Exécutions de programmes parallèles à passage de messages sur grille de calcul*. Habilitation à diriger des recherches de l'université Henri Poincaré, Nancy. Décembre 2009. Rapporteurs : C. Cérin (Paris 13), F. Desprez (INRIA Rhône-Alpes), T. Priol (INRIA Bretagne-Atlantique).
- [2] **Stéphane Genaud**. *Transformations de programmes PEI : applications au parallélisme de données*. Thèse de doctorat de l'université Louis Pasteur, Strasbourg, Janvier 1997. Rapporteurs : Luc Bougé et Patrice Quinton.

Chapitre de livre

- [3] **Stéphane Genaud** et Choopan Rattanapoka. *A Peer-to-Peer Framework for Message Passing Parallel Programs*. Parallel Programming and Applications in Grid, P2P and Network-based System, in *Advances In Parallel Computing Series*. Editor Prof. Dr. Gerhard R. Joubert. IOS Press, juin 2009.

Articles en revues internationales

- [4] **Stéphane Genaud**, Emmanuel Jeannot et Choopan Rattanapoka. Fault-Management in P2P-MPI. *International Journal of Parallel Programming*, Springer, 37(5) :433–461, août 2009.
- [5] **Stéphane Genaud**, Pierre Gançarski, Guillaume Latu, Alexandre Blansché, Choopan Rattanapoka et Damien Vouriot. Exploitation of a parallel clustering algorithm on commodity hardware with P2P-MPI. *The Journal of SuperComputing*, Springer, 43(1) :21–41, jan. 2008.
- [6] **Stéphane Genaud** et Choopan Rattanapoka. P2P-MPI : A Peer-to-Peer Framework for Robust Execution of Message Passing Parallel Programs on Grids. *Journal of Grid Computing*, Springer, 5(1) :27–42, mai 2007.
- [7] **Stéphane Genaud**, Arnaud Giersch, et Frédéric Vivien. Load-balancing scatter operations for Grid computing. *Parallel Computing*, Elsevier, 30(8) :923–946, août 2004.
- [8] Marc Grunberg, **Stéphane Genaud** et Catherine Mongenet. Seismic ray-tracing and Earth mesh modeling on various parallel architectures. *The Journal of Supercomputing*, Kluwer, 29(1) :27–44, juillet 2004.

Articles en revues nationales

- [9] **Stéphane Genaud** et Marc Grunberg. Calcul de rais en tomographie sismique : exploitation sur la grille. *Technique et Science Informatiques*, numéro spécial Renpar, Hermès-Lavoisier, 24(5), pages 591–608, décembre 2005.
- [10] **Stéphane Genaud**. Transformations d'énoncés PEI. *Technique et Science Informatiques*, 15(5), pages 601–618, Hermès, avril 1996.
- [11] **Stéphane Genaud** et Guy-René Perrin. Une expérience d'implantation d'un algorithme systolique sur hypercube. *La Lettre du Transputer et des calculateurs parallèles*, (17), mars 1993.

Conférences internationales avec actes et comité de lecture

- [12] **Stéphane Genaud** et Julien Gossa, Cost-wait Trade-offs in Client-side Resource Provisioning with Elastic Clouds. *4th IEEE International Conference on Cloud Computing (CLOUD 2011)*, juillet 2011. (*papiers acceptés/soumis :36/198, taux : 18%*)

- [13] Pierre-Nicolas Clauss, Mark Stillwell, **Stéphane Genaud**, Frédéric Suter, Henri Casanova and Martin Quinson. Single Node On-Line Simulation of MPI Applications with SMPI. *25th IEEE International Parallel & Distributed Processing Symposium (IPDPS 2011)*, mai 2011. (*papiers acceptés/soumis :112/571, taux : 19%*)
- [14] Virginie Galtier, **Stéphane Genaud** et Stéphane Vialle. Implementation of the AdaBoost Algorithm for Large Scale Distributed Environments : Comparing JavaSpace and MPJ. *International Conference on Parallel and Distributed Systems*, IEEE, déc. 2009. (*papiers acceptés/soumis :91/305, taux : 29%*)
- [15] **Stéphane Genaud** and Choopan Rattanapoka. Evaluation of Replication and Fault Detection in P2P-MPI. *6th IEEE International Workshop on Grid Computing (HPGC), IPDPS 2009*, mai 2009. (*Papier invité*).
- [16] **Stéphane Genaud** and Choopan Rattanapoka. Large-Scale Experiment of Co-allocation Strategies for Peer-to-Peer Supercomputing in P2P-MPI, *5th IEEE International Workshop on Grid Computing (HPGC), IPDPS 2008*, avril 2008.
- [17] Ludovic Hablot and Olivier Glück and Jean-Christophe Mignot and **Stéphane Genaud** and Pascale Vicat-Blanc Primet. Comparison and tuning of MPI implementation in a grid context. *Proceedings of 2007 IEEE International Conference on Cluster Computing (CLUSTER)*, 458–463, september 2007. (*papiers acceptés/soumis :42/106, taux : 39%*)
- [18] **Stéphane Genaud** et Choopan Rattanapoka. Fault Management in P2P-MPI. *International Conference on Grid and Pervasive Computing, (GPC 2007)*, LNCS, Springer, mai 2007. (*papiers acceptés/soumis :56/217, taux : 25%*)
- [19] **Stéphane Genaud**, Marc Grunberg et Catherine Mongenet. Experiments in running a scientific MPI Application on GRID’5000. distingué par le INTEL *best paper award*. *4th IEEE International Workshop on Grid Computing (HPGC), IPDPS 2007*, mars 2007.
- [20] **Stéphane Genaud** et Choopan Rattanapoka. A Peer-to-peer Framework for Robust Execution of Message Passing Parallel Programs. In *EuroPVM/MPI 2005*, LNCS 3666, Springer-Verlag, pages 276–284, septembre 2005. (*papiers acceptés/soumis :61/126, taux : 48%*)
- [21] Marc Grunberg, **Stéphane Genaud**, et Catherine Mongenet. Parallel adaptive mesh coarsening for seismic tomography. In *SBAC-PAD 2004, 16th Symposium on Computer Architecture and High Performance Computing*. IEEE Computer Society Press, octobre 2004. (*papiers acceptés/soumis :32/93, taux : 34%*)
- [22] **Stéphane Genaud**, Arnaud Giersch, et Frédéric Vivien. Load-balancing scatter operations for Grid computing. In *Proceedings of 12th Heterogeneous Computing Workshop (HCW), IPDPS 2003*. IEEE Computer Society Press, avril 2003.
- [23] Romaric David, **Stéphane Genaud**, Arnaud Giersch, Éric Violard, et Benjamin Schwarz. Source-code transformations strategies to load-balance Grid applications. In *International Conference on Grid Computing - GRID’2002*, LNCS 2536, pages 82–87. Springer-Verlag, novembre 2002.
- [24] Marc Grunberg, **Stéphane Genaud**, et Catherine Mongenet. Parallel seismic ray-tracing in a global Earth mesh. In *Proceedings of the 2002 Parallel and Distributed Processing Techniques and Applications (PDP-TA’02)*, pages 1151–1157, juin 2002.
- [25] Eric Violard, **Stéphane Genaud** et Guy-René Perrin. Refinement of data-parallel programs in pei. In *IFIP Working Conference on Algorithmic Language and Calculi*, R. Bird and L. Meertens editors, Chapman & Hall Ed., février 1997. 25 pages.
- [26] **Stéphane Genaud**, Eric Violard, et Guy-René Perrin. Transformation techniques in PEI. In P. Magnusson S. Haridi, K. Ali, editor, *Europar’95*, LNCS 966, pages 131–142. Springer-Verlag, août 1995. (*papiers acceptés/soumis :50/180, taux : 27%*)

Conférences nationales avec actes et comité de lecture

- [27] Marc Grunberg et **Stéphane Genaud**. Calcul de rais en tomographie sismique : exploitation sur la grille. In *Renpar2003*, pages 179–186. INRIA, octobre 2003.

- [28] **Stéphane Genaud**. Techniques de transformations d'énoncés PEI pour la production de programmes data-parallèles. In *RenPar 7*, mai 1995, Mons, Belgique.
- [29] Guy-René Perrin, Eric Violard et **Stéphane Genaud**. PEI : a theoretical framework for data-parallel programming. In *Workshop on Data-Parallel Languages and Compilers*, Lille, mai 1994.

Autres communications

- [30] Christine Carapito, Jérôme Pansanel, Patrick Guterl, Alexandre Burel, Fabrice Bertile, **Stéphane Genaud**, Alain Van Dorsselaer, Christelle Roy. Une suite logicielle pour la protéomique interfacée sur une grille de calcul. Utilisation d'algorithmes libres pour l'identification MS/MS, le séquençage de novo et l'annotation fonctionnelle. Rencontres Scientifiques France Grilles 2011, Lyon.
- [31] Alain Ketterlin, **Stéphane Genaud**, Matthieu Kuhn. Loop-Nest Recognition for the Extraction of Communication Patterns and the Compression of Message-Passing Parallel Traces. Research Report ICPS 11-01. Université de Strasbourg. déc. 2011.
- [32] A. Schaaff, F. Bonnarel, J.-J. Claudon, R. David, **S. Genaud**, M. Louys, C. Pestel and C. Wolf, Work around distributed image processing and workflow management, poster à ADASS 2005, Madrid.
- [33] Marc Grunberg, **Stéphane Genaud**, et Michel Granet. Geographical ISC data characterization with parallel ray-tracing. In *Eos Trans. AGU*, 84(46), *Fall-Meeting Suppl.*, *Abstract S31E-0793*, décembre 2003.
- [34] **Stéphane Genaud**. Applications parallèles sur la grille : mieux vaut il être rapide ou résistant ? *Actes de GridUse-2004*, Workshop "What we have learned", conférence invitée. Supélec Metz, juin 2004.

En cours de soumission

- [35] **Stéphane Genaud**, Julien Gossa et Etienne Michon. Provisioning Cloud resources on the client-side : a cost-performance trade-off approach. IEEE IT Professional Magazine on Cloud Computing. Nov 2011
- [36] Olivier Beaumont, Laurent Bobelin, Henir Casanova, Pierre-Nicolas Clauss, Bruno Donassolo, Lionel Eyraud-Dubois, **Stéphane Genaud**, Sacha Hunold, Arnaud Legrand, Martin Quinson. Towards Scalable, Accurate, and Usable Simulations of Distributed Applications and Systems, ACM Transactions on Modeling and Computer Simulation. Oct 2011.
- [37] Etienne Michon, Julien Gossa, **Stéphane Genaud**. Free elasticity and free CPU power on IaaS Clouds : Promises and Study. Europar 2012. Fev 2012.

2.4 Activités scientifiques

2.4.1 Niveau international

Membre des comités de programmes des conférences internationales :

- 14th IEEE International Conference on High Performance Computing and Communications (HPCC 2012), 2012 (Liverpool, England),
- 13th IEEE International Conference on High Performance Computing and Communications (HPCC 2011), 2011 (Banff, Canada),
- IEEE/ACM International Conference on Grid Computing (GRID'10), 2010 (Bruxelles, Belgique),
- 20th IASTED International Conference on Parallel and Distributed Computing and Systems, 2010, (Marina Del Rey, USA),
- IEEE/ACM International Conference on Grid Computing (GRID'08), 2008 (Tsukuba, Japon),
- International Symposium on Grid and Distributed Computing, 2008 (Hainan Island, Chine),

Relecteur pour de nombreuses revues ou conférences internationales : IEEE Trans. on Distr. and Parallel Systems, J. of SuperComputing, J. of Grid Computing, IEEE Conference on Grid Computing, IEEE CCGrid conference, Europar, IEEE IPDPS conference, ...

2.4.2 Niveau national

- Obtention de la prime d'excellence scientifique (PES) à partir d'octobre 2009.
- membre du comité de rédaction de la revue Technique et Science Informatiques (2005–2009).

Projets en cours

- **Porteur local** pour le projet ANR SONGS (ANR 11 INFR 013-03) (taux déclaré 40%) coordonné par Martin Quinson, LORIA, Nancy (2012-2015) poursuivant le projet Uss-SimGrid (voir ci-dessous). Le projet vise à affiner les objets modélisés pour la simulation (processeurs multi-cœurs, mémoire) ou en ajouter (disque, réseaux spécialisés comme Infiniband) et à fournir des interfaces adaptées à la représentation de systèmes complexes comme des machines HPC ou des Clouds. Je suis responsable du work package sur les clouds.
- Participant au projet blanc ANR E2T2 (ANR 11 SIMI 9) (taux déclaré 15%) coordonné par Peter Beyer, laboratoire PIIM, Université de Provence (2011-2014). L'objectif du projet est d'améliorer la modélisation physique des plasmas de bord dans un tokamak. Dans ce projet, ma tâche est de co-encadrer un doctorant, Matthieu Kuhn avec Guillaume Latu et Nicolas Crouseille (IRMA) pour paralléliser les codes développés par le CEA Cadarache (IRFM) et les physiciens du PIIM.
- **Co-animateur** d'une action d'animation scientifique dans le cadre de l'action de développement technologique (ADT) de Aladdin de l'INRIA, visant à pérenniser l'outil scientifique Grid5000. (07/2008–06/2012). Conjointement à l'ADT, l'animation scientifique est organisée autour de neuf actions d'animations baptisées *défis*. Je co-anime avec Nouredine Melab (LIFL, Université de Lille) le défi "*scalable application for large scale systems (algorithm, programming, execution models)*".

Projets passés

- Participant (taux déclaré 20%) au projet ANR USS-SimGrid (ANR 08 SEGI 022) coordonné par Martin Quinson, LORIA, Nancy (2009 – 2011). Ce projet a été labellisé projet *phare* par l'ANR. L'objectif général du projet était d'élargir les capacités de l'environnement de simulation SimGrid pour satisfaire des besoins plus divers, comme la simulation de systèmes pair-à-pair ou d'environnements de calcul intensif. Mes tâches ont concerné l'enregistrement des traces d'exécutions (des programmes MPI en particulier) afin de les rejouer dans le simulateur. J'ai redémarré le travail commencé à l'université de Hawaï sur l'interface SMPI, qui permet de simuler des programmes MPI sans modification des codes sources. Elle est maintenant fonctionnelle depuis la version 3.5 de SimGrid.
- Participant (taux déclaré 20%) au projet SPADES (ANR 08 SEGI 025) coordonné par Eddy Caron, LIP-ENS Lyon (2009 – 2011). L'objectif était de concevoir et construire un intergiciel capable de gérer un environnement dans lequel la disponibilité des ressources change très rapidement. En particulier, cet intergiciel doit donner accès de manière fugace à des équipements de calculs très haute performance. Mes tâches ont concerné la conception et l'évaluation de l'ordonnanceur travaillant en collaboration avec un système pair-à-pair utilisé pour recenser dynamiquement les ressources disponibles.
- Participant (10%) au projet Masse de Données Astronomiques (ACI Masse de données) coordonné par Françoise Genova, observatoire de Strasbourg (2004 – 2006).
- **Porteur d'un projet d'Action Concertée Incitative**. Projet TAG, pluridisciplinaire dans de l'ACI GRID (Globalisation des ressources informatiques et des données) du ministère de la recherche. Doté d'un budget de 182 K€ et d'un poste d'ingénieur). (12/2001 – 12/2003).

2.4.3 Niveau local

- **Responsable du thème** *programmation parallèle sur les grilles* au sein l'équipe ICPS, du laboratoire LSIIT. Ce thème a compté parmi ses membres : Guillaume Latu (MC), Eric Violard (MC), Romaric David (IR), Benjamin Schwarz (IE en CDD), Arnaud Giersch (doctorant), Choopan Rattanapoka (doctorant) sur la période 2002 à 2007.
- Membre du conseil scientifique du département *Expertise pour la recherche de l'UdS* (sept 2010–). Le comité comprend 17 membres nommés, représentants les équipes scientifiques les plus impliquées par rapport aux équipements de calcul de l'Université. Le rôle du comité est de piloter l'investissement en matière de calcul, et de promouvoir les projets présentant le plus d'intérêt scientifique par attribution de ressources.
- Correspondant depuis 2002 pour l'équipe ICPS auprès du groupe RGE (Réseau Grand Est), action géographique regroupant 9 sites du GDR *Architecture, Système et Réseaux* CNRS (GDR 725 ASR). RGE organise trois fois par an, une journée consacrée à des exposés scientifiques des équipes et à une conférence par un industriel invité.

2.4.4 Jurys de thèse

- Rapporteur de la thèse de Sébastien Miquée, Univ. Franche-Comté (soutenance jan. 2012), *Exécution d'applications parallèles en environnements hétérogènes et volatils : déploiement et virtualisation*, rapporteur C. Cérin (U. Paris 13), encadrants R. Couturier et D. Laiymani (U. Franche-Comté)
- Rapporteur de la thèse de Fabrice Dupros, Univ. Bordeaux 1 (soutenance déc. 2010), *Contribution à la modélisation numérique de la propagation des ondes sismiques sur architectures multicœurs et hiérarchiques*, rapporteur S. Lanteri (INRIA Sophia-Antipolis), encadrants D. Komatitsch (U. Pau) et J. Roman (Institut Polytechnique de Bordeaux).
- Examineur de la thèse d'Heithem Abbès (soutenance déc. 2009), *Approches de décentralisation de la gestion des ressources dans les Grilles*, rapporteurs Mohamed Jmaiel (Université de Sfax) et Franck Capello (INRIA-U. Urbana-Champaign), encadrants Christophe Cérin (U. Paris 13) et Mohamed Jemni (École Supérieure des Sciences et Techniques de Tunis).

2.5 Encadrements

2.5.1 Thèses

1. 10/2011– : encadrement d'Etienne Michon. Taux d'encadrement : 50%, avec Julien Gossa. Financement DGA. La thèse porte sur les problématiques d'allocation de ressources de cloud côté client.
2. 02/2011– : encadrement de Matthieu Kuhn. Financement ANR E2T2. Taux d'encadrement prévisionnel : 20%. Co-encadrants Guillaume Latu pour l'informatique, Nicolas Crouseille (HDR) pour les mathématiques appliquées. La thèse porte sur la parallélisation de modèles numériques pour la simulation de plasmas de bord.
3. 2004–2008 : encadrement de Choopan Rattanapoka. Taux d'encadrement : 100%. Directeur de thèse : Catherine Mongenet. La thèse soutenue en avril 2008 est intitulée *P2P-MPI : A Fault-tolerant Message Passing Interface Implementation for Grids* - rapporteurs : Franck Cappello (INRIA, Orsay) et Thilo Kielmann (Vrije Universiteit, Amsterdam). Choopan Rattanapoka a aujourd'hui un poste permanent d'assistant professor au Department of Electronics Engineering Technology du King Mongkut's University of Technology, à Bangkok (Thaïlande). Publications associées : [20, 6, 18, 5, 16, 15, 4, 3]

4. 2001–2004 : co-encadrement d’Arnaud Giersch avec Frédéric Vivien. Taux de co-encadrement : $\sim 40\%$. Directeur de thèse Guy-René Perrin (ULP). La thèse soutenue en décembre 2004 est intitulée *Ordonnement sur plates-formes hétérogènes de tâches partageant des données* - rapporteurs : Denis Trystram (INPG, Grenoble) et Henri Casanova (UCSD, San Diego). Arnaud Giersch a aujourd’hui un poste de maître de conférences à l’IUT d’informatique de Belfort, université de Franche-Comté. Publications associées : [23, 22, 7]
5. 2000–2006 : co-encadrement de Marc Grunberg avec Catherine Mongenet (inscrit en thèse parallèlement à sa fonction d’ingénieur d’études au Réseau National de Surveillance Sismique). Taux de co-encadrement : $\sim 70\%$. Directeurs de thèse : Catherine Mongenet et Michel Granet (Physicien, ULP). La thèse soutenue en septembre 2006 est intitulée *Conception d’une méthode de maillage 3D parallèle pour la construction d’un modèle de Terre réaliste par la tomographie sismique* - rapporteurs : Thierry Priol (IRISA, Rennes) et Denis Trystram (INPG, Grenoble). Marc Grunberg occupe toujours aujourd’hui un poste d’ingénieur d’études au Réseau National de Surveillance Sismique, École et Observatoire de Géophysique du Globe. Publications associées : [24, 27, 33, 8, 21, 9, 19].

2.5.2 Stages de DEA/Master

1. 2011 : Etienne Michon. Co-encadrement avec Julien Gossa. Mémoire intitulé *Allocation de ressources et ordonnancement côté client dans un environnement de Clouds*, soutenu 06/2011.
2. 2006 : Constantinos Makassikis. Co-encadrement avec Jean-Jacques Pansiot et Guillaume Latu. Mémoire intitulé *Modèle de coût des communications TCP à un niveau applicatif*, soutenu 06/2006.
3. 2006 : Ghazi Bouabene. Mémoire intitulé *Sélection de pairs et allocation de tâches dans P2P-MPI*, soutenu 06/2006.
4. 2004 : Choopan Rattanaopoka. Mémoire intitulé *P2P-MPI : A Peer-to-Peer Framework for Robust Execution of Message Passing Parallel Programs on Grids*, soutenu 07/2004.
5. 2002 : Dominique Stehly. Mémoire intitulé *Ordonnement d’applications parallèles sur la grille*, soutenu 07/2002.

2.5.3 Autres

Encadrement internship INRIA

- *Data Management in P2P-MPI*, Jagdish Acharya, B-Tech de LNMIIT Jaipur, Inde. 3 mois, mai-août 2009.
- *Optimisation de l’opération collective MPI all-to-all*, Antonio Grassi, Master de Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay. 4 mois, avril-août 2008, co-encadré avec Emmanuel Jeannot (AlGorille, LORIA).

Encadrement stage ENSIIE

- *Experimentation de cluster virtualisé avec Nimbus*, Marien Ritzenthaler, ENSIIE 1A. 2,5 mois, juin-août 2010.

Encadrements stages ou projets tutorés d’étudiants de l’UFR d’informatique, université de Strasbourg (150h, un mois plein). Parmi les plus récents :

- *Interfaçage d’un batch scheduler un système de gestion de cloud IaaS*, Vincent Kerbache, TER, 2012, co-encadré avec J. Gossa.
- *Mise en œuvre d’une technique de segmentation par ligne de partage des eaux dans un environnement distribué hétérogène*, Lionel Ketterer, projet tutoré, fév. 2007, co-encadré avec Sébastien Lefèvre (LSIIT).
- *Distribution de calculs de Pricing d’options au modèle Européen sur grille*, Nabil Michraf et Khalid Souissi, projet tutoré, fév. 2006, co-encadré avec Stéphane Vialle (Supelec).
- *Parallélisation de la méthode Adaboost*, Abdelaziz Gacemi, projet tutoré, fév. 2007.
- *Réalisation d’un portail web pour P2P-MPI avec SOAP*, David Michea, projet tutoré, fév. 2005.

- *Parallelisation de la méthode MACLAW*, stage master, avril-juillet 2006, Damien Vouriot, co-encadré avec Pierre Gancarski (LSIIT).
- *Heuristiques d'ordonnancement basées sur des traces d'exécution pour programmes parallèles*, Ghazi Bouabene, stage master, juillet-septembre 2006.
- *Outil de visualisation du réseau P2P dans P2P-MPI*, Ghazi Bouabene, stage licence, juin-août 2005.

2.6 Synthèse quantitative

— Synthèse des activités de publication —

5	articles dans des revues internationales
3	articles dans des revues nationales
1	chapitre de livre dans une publication internationale
15	articles dans des conférences internationales à comité de lecture
3	articles dans des conférences nationales à comité de lecture
5	participations à des comités de programmes de conférences internationales

— Synthèse des activités d'encadrement —

3	thèses soutenues encadrées ou co-encadrées
1,5	thèse en cours co-encadrée
5	stages de DEA et master recherche encadrés ou co-encadrés
7	projets tutorés de niveau master encadrés ou co-encadrés
1	master en apprentissage encadré

3 Activités d’enseignement

Je décris d’abord le contexte dans le lequel se sont déroulées mes activités pédagogiques, puis je présente un tableau récapitulatif des cours assurés, suivi d’une description des contenus de cours les plus récents. J’indique ensuite les principales charges d’intérêt collectif exercées, et je termine par mes motivations en matière de pédagogie.

3.1 Contexte

J’ai été recruté en septembre 1998 sur un poste de maître de conférences en informatique (27^e section) dans une composante de l’Université, l’IECS¹³ dédiée aux sciences de gestion (6^e section). Ce poste préexistant marquait la volonté de la composante de développer les enseignements en systèmes d’information. À mon arrivée, j’ai pris en charge les cours définis qui avaient trait à l’informatique. Puis en 2001, j’ai pris la responsabilité de la filière *systèmes d’information* pour y développer et organiser les enseignements. Mon rôle était de définir et de coordonner les enseignements de la filière, nécessitant en plus des enseignants permanents, de recruter une dizaine de vacataires par an, professionnels de l’industrie dans leur très grande majorité. D’autre part, il m’incombait le suivi des étudiants au travers de la définition de sujets de mémoires de fin d’études, la validation de leur cursus à l’étranger, et l’information vis-à-vis des débouchés professionnels (réunions d’information avec des professionnels, animation d’échanges avec des anciens de la filière). Cette activité s’est arrêtée en 2007 lorsque je suis parti en détachement à l’INRIA.

Les principaux cours que j’ai enseignés dans cet établissement sont : technologies des systèmes d’information, conception des systèmes d’information, gestion de projet et outils pour la gestion de projet, introduction à l’algorithmique et la programmation, architecture des applications web, échange de données informatisées, réseaux. De retour de détachement, j’ai réintégré mon poste tout en effectuant plus d’un tiers de mon service à l’extérieur de l’établissement (procédure facilitée par la fusion des universités strasbourgeoises), à l’UFR de mathématiques et informatique, et dans l’école d’ingénieur ENSIIE¹⁴.

3.2 Enseignements en informatique

- **IUT Informatique** Avant mon recrutement comme maître de conférences, j’ai été pendant deux ans ATER temps plein au département informatique de l’IUT de l’Université Robert Schuman de Strasbourg. J’y ai effectué des enseignements en **algorithmique et programmation** et **système d’exploitation**. Durant cette période j’ai également continué d’assurer des cours du soir au CNAM (un tiers de l’unité de valeur **conception fondamentale des algorithmes** avec Franco Zaroli, de 1996 à 2000).
- **UFR Mathématiques et Informatique** Par ailleurs, j’ai gardé un contact permanent avec le département informatique de l’UFR de mathématiques-informatique de l’Université Louis Pasteur. J’ai assuré des vacations en licence (TD du cours **système d’exploitation**, 24h) et un cours de **systèmes distribués** en DESS (14h par an, de 2001 à 2005). J’ai poursuivi ce cours renommé **applications distribuées, parallélisme et grilles** lors du passage aux masters (36h, effectué pour moitié avec Stéphane Vialle, SUPELEC). Ma participation aux activités du département informatique se manifeste également par l’encadrement de certains Travaux d’Enseignement et Recherche ou d’encadrement de quelques projet 150h (voir liste page 19).

13. Institut d’Enseignement Commercial Supérieur, devenue en 2007, après fusion avec l’IAE, l’École de Management Strasbourg. <http://www.em-strasbourg.eu/>

14. Ecole Nationale Supérieure d’Informatique pour l’Industrie et l’Entreprise, basée à Evry, ayant ouvert un campus à Strasbourg. <http://www.ensiie.fr/>

3.3 Tableau récapitulatif

Année	Filière	Matière	cours	TD	TP	occurrences
93-96 (3)	EOST 3	progr. parallèle de méthode numériques pour la résolution de systèmes linéaires	12			2
	CNAM cycle B	UV conception fondamentale des algorithmes	12	12		1
96-98 (2)	IUT 1 + AS info	algorithmique		24	64	2
	IUT 1 + AS info	C et C++		24	64	2
	IUT 1 info	Système d'Exploitation, Unix	6	16		2
	CNAM cycle B info	UV conception fondamentale des algorithmes	12	12		2
<i>Recrutement Maitre de conférences</i>						
98-05 (7)	Licence info	Système d'exploitation		24		2
	DESS info	Systèmes Distribués	14			4
	DEA info	Modèles de programmation et grille (option)	6			2
	CNAM cycle B	UV conception fondamentale des algorithmes	12	12		2
	ENSPS 2	Outils pour la gestion de projet	6			1
	IECS 2+3	Gestion de projet	20	4		7
	IECS 1	Introduction aux systèmes d'information	18			5
	IECS 2+3	Introduction à la programmation	12	12		7
	IECS 2+3	Technologies internet	24		24	7
	DESS ComElec	Réseaux et échanges de données informatisés	12			4
05-08 (3)	Master 2 info	Applications distribuées, parallélisme et grilles	18			3
	Master Recherche info	Modèles de programmation et grille (option)	6			2
<i>Détachement INRIA</i>						
09-11 (3)	Master 2 info	Applications distribuées, parallélisme et grilles	45			3
	Master 2 info	Fouille de données réparties	4			1
	ENSIIE 1	Systèmes Informatiques	20	10.5	28	2
	EM 2+3	Outils pour la gestion de projet	40	8		2
	EM 2+3	Conception et réalisation de systèmes d'information	12			1
	EM 2+3	Technologies pour les applications web	48		24	2

Les abréviations sont les suivantes :

AS	Année Spéciale : cycle menant à l'obtention du DUT informatique pour étudiants diplômés au moins d'un bac+2 non informatique. Aujourd'hui remplacé par la licence pro QCI.
ComElec	Master Commerce Electronique : M2 en formation continue de l'IECS, http://www.em-strasbourg.eu/docs/Master_CE.pdf
EOST	École et Observatoire des Sciences de la Terre : École d'ingénieur dépendant de l'Université, http://eost.u-strasbg.fr/ecoleing.php
ENSIIE	École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise, campus de Strasbourg. http://www.ensiie.fr/
ENSPS	École Nationale Supérieure de Physique de Strasbourg, http://www-ensps.u-strasbg.fr/
EM	École de Management Strasbourg.

3.4 Contenus des enseignements

Ecole de Management - Systèmes d'information

- **Conception des systèmes d'information** : l'objectif du cours est de faire comprendre aux étudiants en systèmes d'information les approches et techniques possibles ou employées pour analyser ou concevoir un système d'information. Dans l'application basée sur des études de cas, l'étudiant doit distinguer les cas d'utilisation, les flux d'information et les processus, et les données à mémoriser dans le système. Les technologies actuelles permettant de le faire sont aussi présentées.
Principes des méthodes de conception des projets logiciels (ex Merise), des cycles de développement, modélisation avec UML : modèle conceptuel des données avec passage du diagramme de classes au modèle relationnel, représentations dynamiques d'un système.
- **Gestion de projet** : le cours présente la méthodologie usuelle de la planification d'un projet et les outils ou méthodes qui sont disponibles indépendamment des spécificités métiers. La mise en pratique se fait sur des études de cas avec un logiciel.
Démarche de gestion projet, principes de d'évaluation du projet, de découpages et d'estimation des tâches, planification PERT, PERT probabiliste, diagrammes Gantt. Application avec Microsoft Project et Primavera pour la gestion collaborative.
- **Introduction à l'algorithmique et à la programmation** : ce cours d'option à destination d'étudiants en systèmes d'information donne les bases de l'algorithmique. Il a pour objectif de montrer les problématiques techniques des équipes de réalisation des systèmes d'information. La pratique se fait avec javascript et donne lieu à un projet.
Structures de contrôle et de données de la programmation impérative, notions de programmation orientée objet. Application avec JavaScript.
- **Architecture des applications web** : l'objectif est de faire comprendre l'importance prise par ce type d'applications dans les systèmes d'information contemporains.
Notion de client et serveur, serveur web et protocole http, serveur applicatif, architecture n-tiers, principes des langages de script, structures de contrôles et de données de PHP, manipulation des données avec SQL, interfaçage PHP/MySQL. Projet de mise en application durant lequel les étudiants construisent un site web PHP/MySQL.

ENSIIE

- **Systèmes Informatiques** : l'objectif est de faire comprendre de nombreux principes des systèmes informatiques en partant du problème de calculabilité jusqu'aux principes présidant les systèmes d'exploitation modernes, y compris leur composante réseau. Cours conçu par Gérard Berthelot, ENSIIE Evry. TD assurés aussi par Alain Ketterlin (2009) et Julien Gossa (2010).
Machines de Turing, automates, systèmes de numération, quantité d'information (Shannon), codages, structures d'un système d'exploitation, gestion des processus et threads, gestion mémoire, gestion du disque, construction d'un exécutable, interpréteur de commandes, architectures réseaux, programmation socket TCP UDP.

Master informatique

- **Applications distribuées, parallélisme et grilles** : le cours expose le paysage actuel des technologies dans le domaine de l'informatique distribuée ou parallèle à travers quelques technologies clés et montre quelles technologies sont les mieux adaptées à tels ou tels besoins. L'objectif est de donner du recul dans ce domaine aux étudiants, qui sont en passe de rejoindre le monde professionnel. Cours assuré pour moitié avec Stéphane Vialle.
Panorama des systèmes distribués, modèles de programmation RPC (illustration avec Java RMI), web services (illustration avec SOAP), modèle de programmation pair-à-pair (illustration avec JXTA), modèle à passage de messages (MPI), modèle pour le traitement de données massives (MapReduce), principes de mutualisation

des ressources et intergiciels de grille (Globus), principe d'externalisation des traitements ou ressources (Software/Infrastructure/Platform as a Service).

- **Modèles de programmation et grilles** : l'objectif de l'option est de présenter aux étudiants se destinant à la recherche, un état de l'art des problématiques du développement d'applications pour les *grilles* ou les *clouds*. *Catégories de grilles et usages, problématiques : modèles de programmation disponibles, impact de l'hétérogénéité sur les performances, équilibrage de charge applicatif, découverte des ressources, volatilités des ressources et tolérance aux pannes.*

3.5 Principales charges d'intérêt collectif

- Directeur délégué aux système d'information à l'EM Strasbourg (sept 2011–).
En charge d'un service de trois analystes-programmeurs. Mon rôle au sein de la direction est de proposer des choix stratégiques pour les services à développer ou acquérir en matière de système d'information. Je planifie également la ré-organisation des services existants de l'établissement. Ces développements se font en concertation avec l'Université (Direction des Usages Numériques et Direction Informatique). Parmi les projets les plus importants : implantation d'un CRM, gestion du processus de recrutement des vacataires, gestion des étudiants à l'étranger.
- Responsable de la filière d'enseignement *système d'information* à l'IECS, Strasbourg (2001–2007).
- Responsable de la mise en place de l'intranet de l'IECS (2000–2003). Participation au développement, essentiellement assuré par Christos Karacostas, pour lequel j'étais l'encadrant de son mémoire de cycle C du CNAM (soutenu en juin 2001).
- **Commissions de spécialistes** section CNU 27 :
 - titulaire à l'Université Louis Pasteur, Strasbourg, 2004–2008,
 - suppléant à l'Université de Franche-Comté, Besançon, 2005–2008,
 - suppléant à l'Université Henri Poincaré, Nancy, 2006–2008.
- 2010– : Membre du **comité d'experts** (9 membres) pour la section 27 de l'Université de Strasbourg.
- Membre des comités de sélection :
 - poste MC 210 UdS Réseaux et Protocole, 2010,
 - poste MC 1207 Université de Franche-Comté, IUT Belfort Montbéliard, 2010.
- Participation à l'animation de la communauté des enseignants-chercheurs en informatique, à travers l'action collective de l'association SPECIF (Société des Personnels Enseignants et Chercheurs en Informatique de France). Membre du conseil d'administration de SPECIF depuis 2010.

3.6 Projet pédagogique

Je connais le département informatique à travers les interventions que j'y mène régulièrement (un cours annuel en master, interventions ponctuelles en licence), depuis une dizaine d'années. D'autre part, beaucoup de mes collègues en recherche y sont permanents et ne manquent pas de me faire partager le quotidien du département. J'ai donc une connaissance empirique, mais bien sûr partielle, de la vie du département et de son évolution sur quelques années. Néanmoins, je garde une vision extérieure, façonnée en grande partie par mon expérience dans un établissement d'enseignement supérieur en sciences de gestion opérant dans un contexte très compétitif.

Je partage avec les collègues membres du département ce constat et ce regret : la licence et les masters devraient former plus d'étudiants, tout en gardant le même niveau d'exigence académique. Localement, on constate que les attentes de l'industrie concernant le nombre d'étudiants formés ne sont pas satisfaites. Un audit mené par la région sur les formations de type ingénieur avait conclu en particulier que le déficit le plus fort concernait les formations en informatique — ce qui a ensuite débouché sur un appel et l'implantation d'un campus à

Strasbourg de L'ENSIIE. Un autre indicateur est le nombre de demandes d'apprentis de la part des entreprises, qui est supérieur à ce que le master en alternance peut accueillir dans sa configuration actuelle.

Face à cette demande, paradoxalement, les étudiants se sont moins orientés vers la licence informatique ces dernières années. Le problème de la baisse des effectifs est posé à toutes les filières scientifiques. Cependant, les formations en informatique de l'UFR ont deux atouts considérables. D'une part, le marché de l'emploi très actif dans le secteur des technologies de la communication peut permettre de développer davantage les relations avec les entreprises et accroître l'attractivité auprès des étudiants. D'autre part, la réputation de l'Université de Strasbourg permet d'envisager une plus large publicité des formations, en France mais aussi à l'étranger.

Je détaille ci-dessous ma perception de ces deux aspects. Je montre que le levier de la relation avec les entreprises est déjà partiellement actionné. Par conséquent, mon projet vise à actionner le second levier à travers **l'internationalisation des formations du département**.

3.6.1 Relations avec le monde professionnel

Des actions sur le moyen terme pourraient jouer en faveur d'une communication plus développée avec le monde professionnel. En particulier, l'outil de CRM dans le nouveau système d'information de l'université (à la rentrée 2013) devrait permettre de gagner en efficacité dans la gestion de contacts, des stages et des diplômés. La fidélisation des diplômés est un effort qui peut être payant et ouvrir la porte des entreprises pour des stages, des contrats de collaborations, thèses CIFRE, etc. Une autre piste est la création de formation en co-tutelle avec des établissements ayant de fortes relations avec le tissu socio-économique comme l'EM Strasbourg.

Cependant, ce rapprochement avec les entreprises a déjà commencé à s'opérer avec le passage du master Ingénierie du Logiciel et des Connaissances (ILC) en alternance à partir de 2007. Ce passage à l'alternance est de nature à créer des relations durables avec les entreprises. Le recrutement en 2011 d'une chargée de communication et des relations entreprise atteste de l'importance de cette activité. De même, le succès de cette formation auprès des étudiants ne se dément pas après plusieurs années. La formation est même passée en alternance sur les deux années du master en 2010. J'ai le sentiment que cette attractivité de l'alternance bénéficiera sur le long terme à l'ensemble des formations, car elle renforce la réputation du département en matière de liens avec le monde professionnel, et donc d'opportunités de carrière pour les étudiants. Cette perception, si elle se répand, pourrait convaincre plus d'étudiants de s'inscrire en licence informatique. En revanche, il existe un risque dans le contexte actuel, que l'alternance déséquilibre l'orientation des étudiants en asséchant les effectifs des autres filières. Il faut donc renforcer et mettre davantage en avant d'autres perspectives pour les formations classiques, notamment celles plus orientées vers la recherche. Une piste est de construire un ou des masters recrutant à l'international.

3.6.2 Internationalisation

Les étudiants étrangers représentent 20%¹⁵ des effectifs de l'Université de Strasbourg. Ces chiffres montrent l'attractivité de l'Université de Strasbourg. Néanmoins, les formations en informatique et le département informatique en particulier, sont très loin de ces niveaux.

J'ai été récemment confronté à la demande d'un étudiant indien que j'avais encadré en stage INRIA. Celui-ci, en passe de terminer son bachelor en Inde, souhaitait faire acte de candidature pour un master en informatique à l'université de Strasbourg. Cet étudiant ne maîtrisant pas le français, j'ai fait une enquête rapide auprès de mes collègues du département informatique, et je n'ai malheureusement pu que le décourager de faire cette démarche. Cet étudiant s'est donc tourné vers des programmes proposant des enseignements en anglais proposés aux Pays-bas et en Allemagne. Les demandes d'étudiants étrangers possédant déjà une formation de premier cycle de bon niveau sont croissantes. Si des formations de type master international pouvaient les accueillir, ils formeraient un vivier de ressources intéressant, notamment pour la formation doctorale ou les équipes de recherche.

Le projet que j'aimerais promouvoir si j'étais recruté à l'UFR serait la construction d'un programme de master international. Mon établissement de rattachement actuel propose un cursus en formation initiale largement ouvert à l'international (plus de 150 universités partenaires). Les étudiants doivent faire au moins une de leurs trois années dans une Université à l'étranger. Réciproquement, l'établissement reçoit de nombreux étudiants étrangers en échange, qui représentent jusqu'à 50% des étudiants présents. Les étudiants sont ainsi confrontés à des cultures et des méthodes de travail différentes, impliquant des remises en question particulièrement enrichissantes.

15. Dont 47% en doctorat, 31% en Master et 15% en licence. Chiffres 2010.

En discutant avec mes collègues, j'ai noté que la plupart d'entre eux se disaient prêts à faire des enseignements en anglais. Dans le domaine de l'informatique, la technicité de la discipline rend le fond plus important que la forme et je crois qu'il n'est pas nécessaire de maîtriser la langue anglaise comme une langue maternelle pour dispenser un bon enseignement.

Trois solutions au moins sont possibles pour un tel projet : i) intervenir dans une formation délocalisée, ii) construire un dossier Erasmus Mundus, iii) créer ou transformer un master existant avec des enseignements en anglais. La première solution est déjà expérimentée par des Universités comme Bordeaux 1 et Paris 6 depuis 2006 dans un master implanté sur site à Ho Chi Minh-Ville au Vietnam, via le Pôle Universitaire Français. Le retour d'expérience semble mitigé : les difficultés logistiques et financières sont conséquentes alors que les étudiants diplômés sont peu nombreux à poursuivre en France en doctorat ou dans l'industrie. La deuxième solution d'un master Erasmus Mundus est la plus prestigieuse mais aussi la plus longue à construire. Un master Erasmus Mundus est une formation conjointe à plusieurs établissements européens, ouverte à des candidats du monde entier, qui doivent passer leur scolarité sur au moins deux des sites partenaires. Elle assure un fort taux de sélectivité par un recrutement au niveau mondial. Cependant, c'est aussi la solution la plus incertaine en raison de la forte sélectivité, et la plus longue car un des critères d'éligibilité est la pré-existence de collaborations entre les établissements partenaires. La troisième solution semble plus simple à mettre en place. C'est ce qu'a fait Grenoble avec la déclinaison de son master 2 recherche en deux parcours¹⁶ enseigné en français, l'autre en anglais. Le parcours en anglais, MoSIG est perçu, y compris par les étudiants français, comme la formation de prestige.

Conclusion

Je serais donc heureux de pouvoir initier ou participer au montage de tels programmes internationaux, en faisant bénéficier le département de mon expérience de l'EM Strasbourg, qui a développé un réseau extrêmement large d'Universités partenaires à l'étranger. Le paysage de l'offre dans l'enseignement supérieur est actuellement en pleine restructuration, en France mais aussi chez nos voisins, avec la formation de réseaux d'excellence dont l'Université de Strasbourg a vocation à faire partie. C'est donc une période d'opportunités qu'il ne faut pas manquer de saisir.

16. <http://ufrima.imag.fr/spip.php?rubrique94>

4 Documents annexes

Le reste du document fournit :

- la copie d’une pièce d’identité,
- la déclaration de candidature imprimée depuis GALAXIE, datée et signée,
- l’attestation de réussite au diplôme d’HDR,
- le rapport de soutenance d’HDR,
- les trois rapports de pré-soutenance,
- La recommandation de Pierre Tellier, directeur des études de l’antenne de Strasbourg de l’ENSIE,
- La recommandation de Monique Rice, directrice des études à l’IECS.
- La recommandation de Babak Mehmanpazir, directeur de l’IECS de 2005 à 2007.
- La recommandation de Jens Gustedt, chef de l’équipe projet INRIA AlGorille.

*Ce document doit être joint au dossier transmis à
l'intention du comité de sélection*

DÉCLARATION DE CANDIDATURE AU RECRUTEMENT SUR UN EMPLOI DE PROFESSEUR DES
UNIVERSITES

(Campagne 2011)

(décret n° 84-431 du 6 juin 1984 modifié)

Authentification : f9d4e4f148587be57c8239e1d69b82c7 (1300797128576)

adressée au chef d'établissement de : UNIVERSITE DE STRASBOURG

Poste(s) : n°2226

Publié au Journal officiel du : 7 octobre 2009

Section(s) C.N.U : 27 (Informatique)

Profil :

Localisation : Illkirch-Graffenstaden

Article 46-1

Chaire : Non

Je soussigné(e) M.

Nom patronymique : GENAUD

Nom d'usage ou marital :

Prénom : STEPHANE

Date et lieu de naissance : 10/03/1969 - FORBACH

Nationalité : Française

Numen : 15S9801671AYH

N° de qualification : 10127207589

Adresse postale et électronique à laquelle seront acheminées toutes les correspondances		
8A ROUTE ROMAINE		
Code postal : 68230	Ville : TURCKHEIM	Pays : FRANCE
Téléphone : 0389273935	Télécopie :	
Adresse électronique : genaud@unistra.fr		

Fonctions et établissement actuel : Enseignant-chercheur (Maître de conférences)-STRASBOURG

Date de création :

25/02/2011 à 12:02

Date de dernière modification :

25/02/2011 à 13:02

Titres universitaires français :

Diplôme le plus récent : HDR

Titre : Exécutions de programmes parallèles à passage de messages sur grille de calcul

Date de soutenance : 08/12/2009

Lieu de la soutenance : UNIVERSITE HENRI POINCARÉ

Mention : Informatique

Directeur : JENS GUSTEDT

Composition du jury : CRISTOPHE CERIN

FREDERIC DESPREZ

CLAUDE GODART

THIERRY PRIOL

PASCAL BOUVRY

Activités en matière d'enseignement :

o FC CNAM cycle B info. UV programmation [96-00]

o ATER IUT Info. Illkirch. Algo et C/C++, syst. exploit [96-98]

o Maitre Conf. IECS: master gestion, filière système information. Algorithmique, technologies web, base de données, réseau, gestion de projet. 1998-2007 puis EM [09-11]

o UFR math-info. Syst. distribués et applications réparties, master 2 [00-11]

o ENSIE : systèmes infos [09-11]

thème de recherche et mots clés :

Parallélisme, Grille, Cloud, Intergiciel.

- grilles de calcul

- étude et conception d'applications parallèles en environnement distribué hétérogène

- conception d'intergiciel pour les environnements distribués hétérogènes

Activités en matière d'administration et autres responsabilités collectives :

- enseignement : responsable de la filière système d'information de l'IECS. 2000-2007.

- recherche : responsable thème "Grilles" de mon équipe. 2003-2011

- membre de commissions spécialistes 27e : 3 mandats : Univ. L. Pasteur, univ. Franche-Comté et univ. H. Poincaré

- membre comité expert UdS 27e, membre de 2 comités sélection MCF.

- membre de jury pour 2 thèses

Autres titres et diplômes :

Thèse doctorat informatique ULP, 1997.

Transformations de programmes Pei : applications au parallélisme de données. Rapp. P. Quinton, L. Bougé. Dir. thèse G-.R. Perrin.

Travaux, ouvrages, articles, réalisations :

Publications:

* sur SimGrid et SMPI : article 2011 à IPDPS

* sur P2P-MPI :

- article 2007 (j. of grid computing)

- article 2009 sur tolérance aux pannes (intl j. of parallel programming)

- deux articles d'application de P2P-MPI

* Code géophysique : article 2004 (j. of supercomputing) et articles sur grilles de calcul

* Grille : article 2004 équilibrage de charge (j. Par. Comp.)

déclare faire acte de candidature sur l'emploi ci-dessus désigné :

Fait à Strasbourg le 24 mars 2011



Chauhan Nagar

day-freedom

1101682001059STEPHANE<<OLIV6903103MO

ATTESTATION DE REUSSITE AU DIPLOME

Le Président de l'Université atteste que

l' Habilitation à Diriger des Recherches Informatique
a été décernée à

Monsieur GENAUD STEPHANE

né le 10 mars 1969 à FORBACH (057)

au titre de l'année universitaire 2009/2010

Date de soutenance : 8 décembre 2009

Etablissement soutenance : Université Henri Poincaré, Nancy I

Jury

: M. CLAUDE GODART, Président du jury, Professeur des Universités
Université Henri Poincaré, Nancy I
M. CHRISTOPHE CERIN, Rapporteur du jury, Professeur des Universités
Université Paris 13
M. FREDERIC DESPREZ, Rapporteur du jury, Directeur de recherche
INRIA
M. THIERRY PRIOL, Rapporteur du jury, Directeur de recherche
INRIA
M. PASCAL BOUVRY, Membre du jury, Professeur des Universités
Université du Luxembourg
M. JENS GUSTEDT, Membre du jury, Directeur de recherche
INRIA

Ecole doctorale : Informatique, Automatique, Electronique, Electrotechnique (077)

Formation doctorale : Informatique

Fait à NANCY, le 9 décembre 2009



N° étudiant : 29005982

Avis important: Il ne peut être délivré qu'un seul exemplaire de cette attestation. Aucun duplicata ne sera fourni.

24 - 30 rue Lionnois | BP 60120
54003 Nancy Cedex | FRANCE
Tél. : +33(0)3 83 68 20 00
Fax : +33(0)3 83 68 21 00

RAPPORT DE SOUTENANCE

Concernant l'Habilitation à Diriger des Recherches

en *Informatique*
Présentée par *Stéphane GENAUD*
Le *08/12/09*

Le jury a apprécié la présentation d'un travail scientifique fondé sur une démarche rigoureuse et l'expérimentation à large échelle d'applications issues de problèmes importants provenant en bonne partie de disciplines hors de l'informatique.

L'exposé très pédagogique a permis d'illustrer cette démarche de façon très claire et accessible. Il a suscité de nombreuses questions auxquelles M. Genaud a répondu de façon bien argumentée, claire et honnête, démontrant ainsi la maîtrise de son domaine de recherche dans son ensemble.

Ce travail a mis en évidence des verrous scientifiques sur lesquels construire un programme de recherche ambitieux dans le domaine de la programmation parallèle et distribuée à large échelle.

En conclusion, le jury propose que M. Stéphane Genaud soit habilité à diriger des recherches.

Président du Jury - Nom, Prénom et signature :

GODART Claude

Membres du Jury - Nom, Prénom et signature :

PRIOZ, Thierry
DEJOURS, Frédéric
Bouvier, Pascal
GUSTEDT, Jens
Christophe Célin

Rapport sur le manuscrit proposé par
Stéphane Géraud en vue de l'obtention de
l'habilitation à diriger des recherches de
l'université Henri Poincaré – Nancy 1

Christophe Cérin

Vendredi 30 octobre 2009

Préambule

Le manuscript proposé est constitué de trois parties principales et d'une introduction sur deux pages qui présente le document. Nous allons partir de la fin du document, c'est à dire des annexes qui contiennent des copies d'articles les plus significatifs et du curriculum vitæ.

Analyse du Curriculum Vitæ

Le candidat est âgé de 40 ans et il est sur le domaine du parallélisme depuis 1992-1993 où il a obtenu un DESS dans cette discipline, puis une thèse à Strasbourg en 1997. Il a été nommé Maître de conférences dans la foulée (1998) et à Strasbourg. Entre 2007 et 2009 il a été détaché en qualité de Chargé de recherche au Loria de Nancy, dans l'équipe projet INRIA AlGorille. Depuis septembre 2009 il est de retour à Strasbourg et en qualité de Maître de conférences. Il a donc été mobile au cours de ces années.

La lecture du document fait apparaître un effort du candidat pour valoriser son travail dans de bonnes revues internationales (Journal of Grid Computing, Journal of SuperComputing, Parallel Computing) ainsi qu'un chapitre de livre. Les publications en conférences internationales sont d'un bon niveau et elles sont clairement le résultat d'une partie du travail d'encadrement de thèses qui de toute évidence a porté ses fruits. On peut compter une vingtaine de publications en anglais, en conférences internationales ou en revues internationales ce qui témoigne d'une activité soutenue. Au total son curriculum fait état de 31 publications. Il a encadré ou co-encadré 3 thèses à ce jour (à 40%, 70% et 100%).

Stéphane Génaud est co-animateur dans Aladdin (projet porté par l'INRIA), il a été porteur d'une ACI et il participe à différentes ANR (les pourcentages ne sont pas donnés). Il a aussi été membre de différentes commissions de spécialistes puisqu'il est Maître de conférences, il a été responsable de filière d'enseignement et il partage sa charge d'enseignement entre trois établissements Strasbourgeois depuis septembre 2009. Il a aussi été vacataire à différents endroits comme au CNAM, à l'école d'ingénieurs ENSPS, en DESS : il a donc une expérience d'enseignement riche, variée et envers divers publics.

Analyse du rapport scientifique

Le rapport est particulièrement équilibré et cohérent et il concerne les grilles de calcul et plus particulièrement la programmation par passage de

messages en vue de tirer les performances vers le haut : il s'agit de travaux en calcul haute performance. Le candidat vient du parallélisme « classique » et il a su mettre à profit son expérience pour investir de nouveaux champs. Deux axes de travail sont explorés et résumés dans le document et après la lecture complète du dossier on peut dire que le candidat maîtrise les grands instruments (« à la Grid'5000 »), que sa démarche part souvent d'applications réelles et qu'il a su développer et traiter des thématiques importantes qui sont les suivantes :

1. Applications de géophysique (chapitre 2) : l'objectif visé est clairement de dépasser la situation où le modèle de programmation des clusters est le passage de messages et la situation où la communication dans les grilles se fait selon un paradigme maître-esclave. Il montre que certaines applications doivent être abordées dans le cadre d'un modèle de programmation plus fin et pour le cas des grilles il continue à vouloir travailler par passage de messages. De nombreux problèmes de la géophysique sont abordés, avec du couplage de code. Il est à noter que le candidat s'est tourné vers des applications en dehors du champ informatique et que ce travail a été réalisé en collaboration avec des géophysiciens. Il s'agit donc d'un travail inter-disciplinaire. Stéphane Génaud replace les discussions dans le cadre d'une pile logicielle à adapter à la grille (Fig 2.2) et il montre comment le travail de parallélisation de la brique ray2mesh s'est mis en place. Puis Stéphane Génaud motive le travail autour d'un maillage adaptatif (les capteurs ne sont pas répartis de manière régulière) et donne un algorithme glouton de construction d'un tel maillage irrégulier. Le candidat discute également au chapitre 3 de l'opération collective all-to-all, présente un modèle pour la grille, tire des conclusions pertinentes à partir d'expériences sur de vraies plateformes (il s'agit ici d'une sous partie autour de l'équilibrage de charge statique et la conclusion est attendue dans le sens où il faudrait l'adapter au cas dynamique). La copie correspondante de l'article est donnée dans les annexes. Comme autre perspective du travail pour la problématique de l'équilibrage de charge, mon avis serait d'examiner la technique de résolution quand les fonctions de coûts ne sont plus affines mais convexes. L'hypothèse « affine » permet de transformer le problème en un problème linéaire que l'on résout par des techniques connues. Par contre quand la fonction de coût (du travail réalisé localement sur un nœud n'est plus linéaire) la littérature sur l'équilibrage des charges n'est pas développée à notre connaissance. Un degré supplémentaire de complexité serait de considérer en plus des communications de coût non-affine ;

2. P2P-MPI et applications gridifiées : cette activité liée aux systèmes pairs à pairs qui à pris de l'importance dans la communauté française depuis 4-5 ans est complète dans le sens où elle introduit un nouvel intergiciel avec toutes les propriétés requises, y compris la tolérance aux fautes. De nombreux articles ont été écrits pour ce travail en profondeur qui discutent à la fois des techniques utilisées dans cet outil mais aussi des applications que l'on peut coder, déployer à large échelle avec l'outil. À la lecture des chapitres qui s'y rapportent on mesure que le candidat a développé une activité en architecture des intergiciels jusqu'aux portages d'applications, en passant par la validation et l'écriture de codes non triviaux parce que faisant intervenir différentes techniques de programmation. C'est donc un travail d'ampleur qui est résumé dans le manuscrit. A chaque fois, le candidat précise ses hypothèses, son cadre de travail, ce qu'il traite et ce qu'il ne traite pas ce qui rend le document particulièrement lisible.
3. Les chapitres 5 et 6 s'intéressent à des applications parallélisées par le candidat dans le cadre d'un travail pluri-disciplinaire. Le chapitre 5 utilise l'expérience du candidat en matière de Java Message Passing (chapitre 4) et vise à comprendre les limites à l'extensibilité de l'application. Le chapitre 6 s'intéresse à une méthode de classification non-supervisée et de nouveau Stéphane Génaud intervient sur la parallélisation de la méthode en mettant en avant ses compétences sur la gestion des communications sur la grille. L'application permet aussi de « valider » le travail du chapitre 4 (autour de P2P-MPI) en concluant que le couplage de celui-ci avec l'application n'est pas difficile à réaliser.

Perspectives de travail

Le chapitre 7 est dédié à la présentation de pistes de travail. Celles-ci sont organisées selon trois axes :

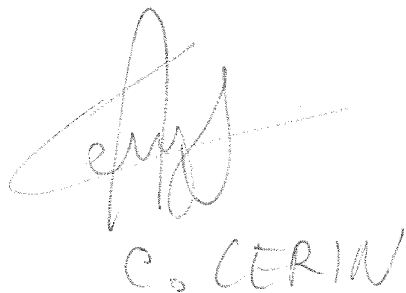
1. Simulation : ce projet correspond à un projet récent qui vise en partie à augmenter les capacités de SimGrid pour l'étendre vers la captation de modèles réalistes de passage de messages avec des communications non-bloquantes ; Stéphane Génaud propose également de s'attaquer à la modélisation des communications TCP, non plus dans le cadre d'un cluster (pour lequel les modèles dérivés de LogP peuvent servir) mais dans le cadre de clusters de clusters. Mon avis serait plutôt de se servir de LogP entre les sites en ajoutant du broadcast intra-sites dans le modèle afin de capter les communications multi-sites en général : on en discutera peut-être à la soutenance ;

2. Gestion de données : il s'agit d'étudier la réplication en prenant des informations auprès de l'application (en supposant par exemple que le code source est disponible) ou encore d'ordonnancer des workflows en fonction des données ;
3. Stratégies d'allocation et de découverte de ressources : ce projet correspondant en partie au projet SPADES dans lequel le candidat est impliqué. Il vise à mieux comprendre et à modéliser des machines à base de processeurs multi-coeurs et de rendre les intergiciels plus réactifs. Pour la partie découverte, Stéphane Génaud propose de se concentrer sur les problèmes de connectivité des ressources d'un sous réseau à un autre. Sa conviction est que « l'intergiciel (du futur) devra masquer la complexité des matériels et guider l'utilisateur dans le choix des ressources. »

Conclusion

Au final, Stéphane Génaud, par son manuscrit, montre ses qualités d'autonomie scientifique dans plusieurs disciplines du calcul haute performance en contexte de grille. Son travail est bien ancré dans des problématiques importantes pour la communauté : couplage de code intelligent, tolérance aux fautes et P2P pour ne citer que quelques exemples. Son parcours professionnel témoigne également d'une volonté à s'ouvrir à des collaborations extérieures à sa discipline d'origine d'une part et au sein d'institutions variées d'autre part. Sa démarche est exemplaire à bien des égards.

La candidature de Stéphane Génaud mérite donc d'être soutenue et pour toutes les raisons données ci-dessus je suis favorable à une autorisation de soutenance devant le jury proposé.



C. CERIN



Laboratoire de l'Informatique du Parallélisme

Unité mixte CNRS - ENS Lyon - INRIA - UCB Lyon n°5668

Rapport sur le document d'HDR de Stéphane Genaud intitulé « Exécution de programmes parallèles à passage de messages sur grille de calcul »

Lyon, le 16/11/09

1. Introduction

Le portage d'applications de calcul scientifique sur les grilles est un problème crucial. Ces plates-formes à hautes performances étaient utilisées au départ soit pour le traitement de tâches parallèles faiblement couplées ou pour l'utilisation en mode batch « classique ». D'autre part l'utilisation des systèmes pair-à-pair s'est souvent limitée à des applications constituées de nombreuses tâches indépendantes (en mode client-serveur). On peut constater enfin qu'en l'absence d'un modèle de programmation spécifique aux grilles, de nombreuses applications écrites en passage de message (comme avec MPI) doivent être portées sur ce type de plates-formes disponibles en production pour les scientifiques et ingénieurs des mondes académiques et industriels.

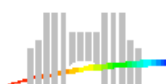
Durant ces travaux de recherche entre le LSIIT de Strasbourg et le Loria à Nancy, Stéphane Genaud s'est intéressé à l'utilisation du passage de message pour le portage d'applications scientifiques sur des plates-formes hétérogènes et distribuées. Que ce soit à travers la conception d'algorithmes d'équilibrage des charges ou le développement d'un intergiciel pour les systèmes pair-à-pair, il s'est attaché à les valider sur des plates-formes réelles à grande et échelle et sur diverses applications aux besoins importants.

2. Contributions scientifiques

Le document de Stéphane Genaud est divisé en deux parties suivies de deux annexes. La première partie décrit le portage d'une application de simulation numérique sur une grille et est divisée en deux chapitres. La seconde partie décrit P2P-MPI, un intergiciel permettant l'écriture d'applications pour les systèmes pair-à-pair en passage de message avec le standard MPI et certaines applications qui l'ont validé. Enfin, une première annexe donne une série de copies d'articles et la seconde présente son cv et son rapport d'activités.

L'application de géophysique étudiée par Stéphane Genaud en collaboration avec l'institut de physique du globe de Strasbourg est typiquement le type d'application pour laquelle le portage sur une grille est un challenge. En effet, plusieurs niveaux de parallélisme sont possibles. Le **chapitre 2** décrit l'application elle-même et la parallélisation de deux de ses noyaux, à savoir le code qui calcule le tracé de rai dans un maillage et celui qui s'occupe du maillage adaptatif. La parallélisation est « classique » mais ses résultats sont excellents, à part du côté de

**Ecole Normale Supérieure de Lyon, 46 Allée d'Italie, 69364 Lyon Cedex 07,
France**



UNIVERSITÉ
DE
LYON

l'extensibilité du maillage adaptatif qui utilise des communications globales. Ces parallélisations montrent clairement l'intérêt d'utiliser des plates-formes hétérogènes et distribuées. Les codes ont été testés sur deux grilles nationales et avec des données de tailles significatives. Le **chapitre 3** présente une étude d'automatisation de l'équilibrage de charge statique pour la distribution du travail d'une application parallèle sur une plate-forme hétérogène. Cette approche innovante est extrêmement intéressante et mériterait de faire partie d'un logiciel de distribution de travail en mode maître-esclaves.

La bibliothèque P2P-MPI décrite dans la seconde partie (et en particulier dans le **chapitre 4**) propose une programmation par passage de messages avec le standard MPJ (MPI pour Java) sur des plates-formes volatiles. Développé en Java et basé sur une base logicielle existante et de nombreux développements originaux, cet intergiciel pair-à-pair apporte plusieurs optimisations importantes comme l'ordonnancement basé sur une heuristique de partitionnement de graphes et la tolérance aux pannes utilisant la réplication active (avec évaluation théorique de la robustesse). Le **chapitre 5** décrit une application d'apprentissage automatique parallélisée à l'aide de MPJ. Une comparaison avec JavaSpace est également donnée tant sur le plan de l'expressivité que des performances. Le **chapitre 6** décrit une application de classification non-supervisée à l'aide de P2P-MPI. Cette parallélisation a montré à la fois les bonnes performances de P2P-MPI mais aussi certaines de ses limitations qui ouvrent des perspectives de recherche intéressantes. Le **chapitre 7** qui conclut cette partie présente une synthèse des résultats et des perspectives de recherche variées et ambitieuses. Les travaux proposés ont pour thème à la fois du côté la modélisation des applications et de la simulation et d'autre part l'ordonnancement de workflows sur des ressources dynamiques.

3. Rapport d'activité

Le cv donné dans l'**annexe B** montre l'activité de recherche et d'enseignement soutenue de Stéphane Genaud. Ses travaux ont été largement publiés dans un chapitre de livre, des journaux et des conférences internationales de renom. Il a également encadré plusieurs thèses (dont une à 100%) et participé à de nombreux projets nationaux.

4. Conclusion

Les travaux de Stéphane Genaud autour de la parallélisation d'applications utilisant le paradigme de programmation à passage de message sur des plates-formes aussi variées que des grappes, des grilles ou des systèmes pair-à-pair sont novateurs et importants pour la communauté du calcul scientifique. En plus de travaux théoriques de qualité, les outils et les applications parallélisées ont été validés en vraie grandeur et sur des plates-formes à grande échelle. Il s'est attaché à comprendre les applications étudiées et à les optimiser pour obtenir les meilleures performances tout en gardant à l'esprit le problème de la programmation sur plates-formes parallèles. Ses perspectives de recherche sont innovantes et extrêmement ambitieuses. Je suis persuadé que les qualités de chercheur aux talents multiples de Stéphane Genaud lui permettront de faire des avancées significatives dans le domaine des intergiciels pour les grilles, les systèmes P2P et les Clouds.

Pour toutes ces raisons, je recommande très chaudement que Stéphane Genaud présente ses travaux pour l'obtention de l'Habilitation à Diriger des Recherches de l'Université Henri Poincaré – Nancy I, spécialité informatique.



Frédéric Desprez
INRIA

Rapport sur le document d'habilitation à diriger des recherches de Stéphane Genaud

Les travaux de recherche de Stéphane Genaud se situent dans le domaine de la programmation et de la conception d'applications adaptées aux grilles informatiques : des infrastructures hétérogènes à la fois parallèles et distribuées. Une des caractéristiques des travaux dirigés par S. Genaud a été l'expérimentation systématique des techniques et des environnements logiciels sur des infrastructures de grilles informatiques à grande échelle et notamment Grid'5000. Le mémoire, rédigé par S. Genaud, est constitué d'une présentation synthétique de ses activités de recherche et se termine par une sélection de papiers les plus significatifs ainsi que d'un CV. J'examinerai ce mémoire en regardant d'abord les activités de recherche pour conclure par une évaluation d'ensemble des activités de S. Genaud.

ACTIVITES DE RECHERCHES

Le mémoire décrivant les travaux de S. Genaud est organisé en deux parties décrivant ses deux principales contributions.

La **première partie** est consacrée à une application de tomographie sismique avec pour objectif la construction d'un modèle de terre réaliste ainsi qu'à sa « gridification ». Les travaux de S. Genaud concernent la parallélisation d'algorithmes en traitant les problèmes de distribution de données, afin de répartir les données dans les mémoires locales des processeurs, et la distribution des calculs pour l'équilibrage de charge. Il a proposé une technique, basée sur un algorithme de type glouton, pour la construction de maillages adaptatifs adaptés à la tomographie sismique dont il a proposé une parallélisation. Ce travail de nature algorithmique est accompagné de plusieurs expérimentations sur des grappes de PC ainsi que sur des grilles de calcul à l'échelle nationale (TAG, Grid'5000). Les résultats obtenus ont permis de montrer l'importance des caractéristiques du réseau utilisé dans les grilles sur les performances globales de l'algorithme. Par exemple la distance, donc la latence, séparant les machines de la grille n'est pas un facteur qui limite les performances de l'algorithme. Au final, les travaux de S. Genaud, menés en étroite collaboration avec l'Institut Physique du Globe à Strasbourg, ont permis de réaliser une tomographie sismique utilisant l'ensemble des données enregistrées par les réseaux de surveillance sismique depuis 1965 ce qui n'avait pas été réalisé auparavant. Ceci a donné lieu à la réalisation de plusieurs logiciels qui ont fait l'objet d'une publication sous une licence GPL.

La **deuxième partie** du document décrit les travaux autour de P2P-MPI ainsi que la parallélisation de deux applications. La conception de l'environnement P2P-MPI est issue du constat que les environnements logiciels autour des grilles, et plus particulièrement Globus et les bibliothèques logicielles associées, sont souvent incomplets et difficiles à maîtriser. Il est souvent impossible de mener à bien des expérimentations dans un environnement hétérogène alors que c'est une caractéristique attendue des grilles informatiques. Pour remédier à cette situation, et fort de son expérience avec les environnements existants, S. Genaud a conçu l'environnement P2P-MPI fondé sur la machine virtuelle Java offrant aux programmeurs une mise en œuvre de MPI en Java. Partant du constat que les pannes sont aussi des éléments importants à prendre en compte dans la conception d'un environnement logiciel pour les grilles, il a proposé un mécanisme de tolérance aux fautes utilisant un protocole de réplication. Il a ensuite mené plusieurs études pour montrer le bien fondé de l'approche notamment sur le taux optimal de réplication. Il s'agit là d'un travail original qui fait l'objet d'une mise à disposition de l'environnement en open source. Les deux autres contributions décrites dans cette partie du document concernent la parallélisation de deux applications dans le domaine de l'apprentissage automatique, appliquée à la détection de visage, et de la classification. Pour la première application, S. Genaud a proposé deux approches : la première fondée sur l'utilisation de JavaSpace (approche à la Linda) et l'autre utilisant P2P-MPI ce qui permet de comparer deux paradigmes de communication : mémoire partagée et échange de messages. S. Genaud a plus particulièrement étudié le problème de l'équilibrage de charge en

déterminant la granularité qui donne les meilleures performances. Pour la deuxième application, la parallélisation proposée s'appuie sur P2P-MPI et montre une très bonne performance pour une configuration d'une soixantaine de processeurs.

Au final, les contributions scientifiques de S. Genaud sont d'excellente qualité. S. Genaud a su mener un programme de recherche pluridisciplinaire en lançant plusieurs collaborations avec des équipes de recherche ayant des applications dont la parallélisation est indispensable pour réduire les temps de traitement. A partir de ces expériences et de sa très grande connaissance des applications, S. Genaud a proposé des techniques d'équilibrage de charges ainsi qu'un environnement logiciel complet adapté aux grilles informatiques. Il est également important de souligner l'effort réalisé pour l'expérimentation de toutes ces techniques en s'appuyant sur la grille de recherche Grid'5000. Les résultats obtenus, notamment pour la tomographie sismique, ont permis de montrer que des grilles pouvaient être des infrastructures distribuées capables d'offrir de très bonne performance. Ces résultats sont souvent repris dans les exposés réalisés par les responsables de Grid'5000 car ils sont exemplaires. Le projet de recherche proposé par S. Genaud contient plusieurs pistes : outils de simulation de programme parallèle à échange de messages, gestion de données, ordonnancement de workflows et déploiement d'applications dans les grilles. Ce projet de recherche est très ambitieux et il y a sans doute lieu de bien définir les priorités en fonction des forces disponibles.

CURRICULUM VITAE

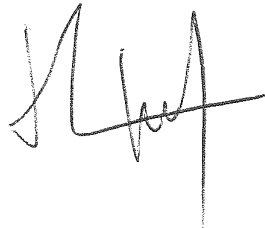
Les annexes du document contiennent un CV ainsi qu'une sélection d'article. Il possède un très bon dossier de publications : **sept** revues internationales et **treize** papiers dans des conférences et workshops internationaux. Un des papiers (HPGC 2007, « *Experiments in running a scientific MPI application on Grid'5000* ») a été distingué (Intel Best Paper Award). En disant cela, je prend bien évidemment aussi en compte que la plupart de ces publications ont nécessité des développements logiciels conséquents et complexes ainsi que l'utilisation de plateformes comme Grid'5000. En ce qui concerne ses activités d'encadrement, j'ai relevé le co-encadrement de trois doctorants qui ont soutenu leurs thèses avec succès. Sur la prise de responsabilités, S. Genaud a été notamment le porteur d'un projet ACI GRID dont l'évaluation a été très positive. Il participe à plusieurs projets ANR ce qui montre une parfaite intégration dans la communauté scientifique nationale.

CONCLUSION

En résumé, les contributions scientifiques de S. Genaud sont nombreuses et d'excellente qualité, ceci est attesté par les publications dans des conférences et journaux internationaux reconnus. Elles combinent de manière équilibrée des aspects conceptuels et expérimentaux. S. Genaud mérite sans aucun doute l'attribution d'une habilitation à diriger des recherches pour l'ensemble de ses contributions, sa capacité à diriger des recherches au regard des encadrements réussis de jeunes doctorants et d'un programme de recherche ambitieux.

Le 10 novembre 2009,

Thierry Priol
Directeur de recherches INRIA



Mesdames et Messieurs les Membres du
Comité de Sélection
Université de Strasbourg
67000 STRASBOURG

Objet : lettre de recommandation

Illkirch, le 15 mars 2011

M. Pierre Tellier
Directeur
ENSIIE Strasbourg
Courriel :
pierre.tellier@ensiie.fr
www.ensiie.fr

Mesdames, Messieurs,

C'est avec un immense plaisir que je sollicite votre attention sur la candidature de Monsieur Stéphane GENAUD au poste de Professeur des Universités en Informatique à l'I.U.T. Robert Schuman d'Illkirch.

J'ai le privilège de compter Stéphane dans l'équipe pédagogique de l'ENSIIE Strasbourg, grande école d'ingénieurs en informatique, créée en 2009.

Durant ces deux années, Stéphane a démontré ses capacités à s'investir dans une formation exigeante, tout en continuant à mener l'activité de recherche qui était la sienne.

La difficulté de son intervention consistait à mettre en place et à dispenser un enseignement dans le domaine des systèmes et réseaux, de haut niveau et répondant à nos exigences de professionnalisation. Dans un contexte difficile de création, Stéphane a toujours été disponible pour partager son expertise et sa rigueur et jouer un rôle moteur dans la qualité et l'organisation de notre formation.

Compte tenu des grandes similitudes de fonctionnement et d'objectifs de nos établissements, Stéphane possède de toute évidence toutes les qualités qui feront de lui un atout pour l'I.U.T. Robert Schuman, et c'est pour cela que je le recommande avec la plus grande énergie et sincérité.

Veuillez recevoir, Mesdames, Messieurs, mes salutations les meilleures



Pierre TELLIER
Directeur



Programme Grande Ecole

Masters universitaires

Executive Education

Strasbourg, le 15 mars 2011

RECOMMANDATION

Stéphane Genaud a intégré l'Ecole de Management Strasbourg en septembre 1998 comme maître de conférences en informatique.

Il est chargé d'assurer des cours d'informatique appliquée à la gestion. En particulier, il enseigne ou a enseigné

- Outils pour la gestion de projet : 24h - Master Grande Ecole (MGE)
- Introduction à la programmation : 24h - MGE 2/3^{ème} année
- Technologies pour les applications web : 48h - MGE 2/3^{ème} année
- Introduction aux systèmes d'information: 18h - MGE 1^{ère} année
- Réseaux et échanges de données informatiques : 16h - Master Commerce électronique (Formation continue)

Les campagnes d'évaluation des enseignements et de la vie à l'école, administrées aux étudiants chaque semestre, ont souligné les compétences pédagogiques de Monsieur Genaud, sa grande disponibilité, ainsi que la qualité des documents proposés.

Monsieur Genaud a dirigé et structuré la filière Management des systèmes d'information du Master Grande Ecole. Dans le cadre de cette mission, il a recruté des enseignants vacataires et permanents (PAST). Sa connaissance du monde de l'entreprise lui a permis de convaincre des spécialistes, non seulement reconnus dans leur domaine mais aussi pédagogues, d'assurer des enseignements dans la filière.

Monsieur Genaud a managé et dynamisé son équipe pédagogique afin de proposer une formation répondant à la fois aux besoins des entreprises et aux attentes des étudiants.

Par ailleurs, Monsieur Genaud s'est impliqué dans la vie de l'Ecole en assurant des charges administratives comme le recrutement de Webmaster, la mise en place de procédures informatiques répondant à des besoins administratifs, en particulier des services de scolarité et relation entreprises.

Depuis septembre 2009 (retour de l'INRIA), nous co-pilotons le développement des systèmes d'information de l'Ecole

En tant que Directrice des Etudes du Master Grande Ecole (fin de mission septembre 2009) et Co-Responsable des SI, je travaille en étroite collaboration avec Monsieur Genaud, et c'est avec plaisir que je tiens à souligner que dans toutes ses activités, ce dernier fait preuve d'un dévouement exemplaire au service public. Sa conscience professionnelle, ses compétences, l'attention qu'il a portée aux problèmes spécifiques de gestion, lui ont valu une grande estime de la part des étudiants et de ses collègues.

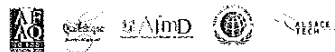
Pour valoir ce que de droit

Monique RICE
Maître de conférences

61, avenue de la Forêt-Noire
F-67085 Strasbourg Cedex
tél.: +33 (0)3 90 41 42 00
fax : +33 (0)3 90 41 42 70
e-mail : contact@em-strasbourg.eu

www.em-strasbourg.eu

EAT Strasbourg Partenaires



Contrat de l'Etat FRS



Programme Grande Ecole

Masters universitaires

Executive Education

Strasbourg, le 22 mars 2011

Lettre de recommandation

Monsieur Stéphane GENAUD a pris la direction de la filière Management des systèmes d'information du Programme Grande Ecole de l'IECS (aujourd'hui Ecole de Management Strasbourg) en 2001 et a occupé cette fonction jusqu'en 2007, date à laquelle il est parti temporairement en détachement.

En tant que responsable de la filière Management des systèmes d'information, Mr Genaud avait la charge de définir le programme pédagogique de cette filière en accord avec la direction de l'école et les exigences du référentiel de la Commission d'Evaluation des Diplômes de Gestion. A ce titre, il avait également la charge de recruter tous les intervenants professionnels et académiques et de coordonner leurs interventions pédagogiques.

Il devait assurer le suivi des étudiants (proposition de thèmes de mémoires de fin d'études, encadrements de certains mémoires, validation des cours suivis par les étudiants à l'étranger).

Enfin, il participe en tant que membre du comité pédagogique bi-mensuel du programme que je présidais pour déployer la maquette pédagogique et d'apporter les réponses aux problèmes de fond ou de forme liés à la scolarité de nos étudiants.

Monsieur Genaud a fait preuve d'une implication exemplaire et a su enrichir ce programme en y concevant de nouveaux modules de cours (développement des cours de management de projet, en e-business, SAP, etc.). Sa contribution a été bénéfique pour nos étudiants et diplômés tout le long de leur cursus.

Mise à part la mission pédagogique extrêmement lourde à porter, Monsieur Genaud s'est impliqué dans de nombreux aspects de la vie de l'école :

Développements d'application sur l'intranet de l'Ecole lors de sa mise en place, encadrement du stage « Ingénieur CNAM » de notre responsable informatique, expertise technique sur les problématiques informatiques quant aux choix de telle ou telle solution.

Je ne peux qu'exprimer notre grande satisfaction quant à notre collaboration avec Monsieur Genaud tant sur le plan professionnel qu'humain.



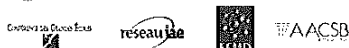
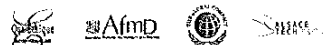
Babak MEHMANPAZIR

Directeur délégué

61, avenue de la Forêt-Noire
F-67085 Strasbourg Cedex
tél.: +33(0)3 68 85 80 00
fax : +33 (0)3 68 85 85 93
e-mail : contact@em-strasbourg.eu

www.em-strasbourg.eu

E.M.Strasbourg
Partenaires





nouveau thème du “Cloud” avec une publication dans une conférence très sélective.

Il a déjà encadré (resp. co-encadré) avec succès trois thèses et plusieurs DEA, masters et internships.

Les perspectives de recherche que Stéphane GENAUD élabore pour les années à venir me semble bien conçues, pertinentes et adéquates :

Entre autres, il est responsable d’un axe sur IaaS (Infrastructure as a Service) dans le projet ANR SONGS. Ce projet est un pas important vers la structuration de la communauté française autour de la thématique de la simulation pour les grilles et le Cloud. Stéphane GENAUD apparaît ici sur des tâches importantes et intéressantes. Il est porteur local dans son laboratoire et leader du workpackage “Cloud”. Le travail sur SMPI va aussi continuer dans ce cadre.

Un deuxième volet des recherches actuels et futures concerne toujours la parallélisation des application. Là, Stéphane GENAUD a de bons contacts aux milieu scientifique (physique de fusion, biologie, chimie) qui donnent lieu à des collaborations intéressantes et permettent de faire le transfert des technologies. Ce volet me semble d’autant plus intéressant que ça fait de lui aussi un élément pivot de la collaboration math-info dans le nouveau LabEx IRMIA de l’université de Strasbourg.

Le contact au jour le jour au sein de mon équipe m’a permis de me faire une très bonne impression des capacités scientifiques et humaines de Stéphane GENAUD. Il est un collègue très compétent, très motivé et plein d’élan, qui réussit bien à transmettre cet enthousiasme aux collègues et aux étudiants stagiaires qui sont sous sa responsabilité.

Ainsi j’éprouve la plus grande estime pour Stéphane GENAUD, son travail scientifique et ses capacités de direction et je le soutiens entièrement et sans aucune réserve pour un poste de professeur. Son recrutement en tant que professeur des université sera un enrichissement pour tous laboratoire et équipe qui voudra bien l’accueillir.

Jens GUSTEDT
Directeur de recherche INRIA
Responsable de l’Équipe-Projet AlGorille
Responsable du département réseaux, systèmes et services du LORIA

5 Copies d'articles

Je joint ci-après six copies d'article. Ils sont classés dans un ordre chronologique inverse :

1. Article sur la simulation de programmes MPI dans le simulateur SIMGRID (SMPI)
Single Node On-Line Simulation of MPI Applications with SMPI.
Pierre-Nicolas Clauss, Mark Stillwell, **Stéphane Genaud**, Frédéric Suter,
25th IEEE International Parallel & Distributed Processing Symposium (IPDPS 2011), IEEE Computer Society Press, mai 2011.
2. Article sur la tolérance aux pannes et la détection de pannes en P2P-MPI.
Fault-Management in P2P-MPI.
Stéphane Genaud, Emmanuel Jeannot et Choopan Rattanapoka.
International Journal of Parallel Programming, Springer, 37(5) :433–461, août 2009.
3. Article sur la parallélisation d'une méthode de clustering avec P2P-MPI.
Exploitation of a parallel clustering algorithm on commodity hardware with P2P-MPI.
Stéphane Genaud, Pierre Gançarski, Guillaume Latu, Alexandre Blansché, Choopan Rattanapoka et Damien Vouriot.
The Journal of SuperComputing, Springer, 43(1) :21–41, jan. 2008.
4. Article sur la conception de P2P-MPI
P2P-MPI : A Peer-to-Peer Framework for Robust Execution of Message Passing Parallel Programs on Grids. Stéphane Genaud et Choopan Rattanapoka.
Journal of Grid Computing, Springer, 5(1) :27–42, mai 2007.
5. Article sur l'équilibrage de charge.
Load-balancing scatter operations for Grid computing.
Stéphane Genaud, Arnaud Giersch, et Frédéric Vivien.
Parallel Computing, Elsevier, 30(8) :923–946, août 2004.
6. Article sur la parallélisation du tracé de rai de l'application de géophysique.
Seismic ray-tracing and Earth mesh modeling on various parallel architectures.
Marc Grunberg, **Stéphane Genaud** et Catherine Mongenet.
The Journal of Supercomputing, Kluwer, 29(1) :27–44, juillet 2004.