

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 important 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

