

Caractérisation globale de l'exécution de jobs

Fabrice Gadaud*, Guillaume Duquesnay*

*Compagnie Générale de Géophysique,
1, rue Léon Migaux
91 341 Massy CEDEX
{fgadaud, gduquesnay}@cgg.com,
<http://www.cgg.com>

Résumé. La *caractérisation globale de l'exécution de jobs* passe par l'exploitation de mesures recueillies sur les machines en production. Afin de répondre à la problématique, il est nécessaire de tenir compte des différents types de données, ainsi que de la dualité de la caractérisation : *statique* et *dynamique*. Une solution technique répondant aux contraintes est proposée. Elle repose sur l'utilisation de *SVM* afin de détecter des phases, et à un niveau supérieur, à un *réseau bayésien* afin d'automatiser l'analyse de *modèles de Markov enrichis*. Ceux-ci sont introduits comme la base formelle et synthétique de description du comportement du job, aussi bien sur un système batch, que parallèle. Enfin, les résultats obtenus à l'aide d'un prototype sont discutés.

Introduction

Dans le cadre de son activité de traitement des données sismiques, le groupe CGG¹ exploite une puissance de calcul importante, répartie en différents points du monde. Les calculs scientifiques effectués sont extrêmement lourds et réalisés depuis plusieurs années sur des machines dédiées, au début des mainframes, puis principalement des clusters de PCs sous GNU/Linux. La CGG est devenue l'entreprise pionnière dans l'exploitation, en production, de ce type d'architecture². Ainsi, toujours dans le souci d'améliorer ses performances, la CGG a décidé d'optimiser ses ressources, et donc de mieux les connaître. Pour cela, elle dispose de données relatives aux jobs³, recueillies dans ses différents centres, sur chaque unité de calcul. C'est donc tout naturellement vers les techniques d'extraction de connaissances que des recherches se sont portées, dans le but de valoriser les données de manière innovante, en dépassant l'approche statistique de base⁴ souvent rencontrée dans ce type de cas.

¹Compagnie Générale de Géophysique

²DELL a remis son prix "centre d'excellence de recherche" à la CGG, en septembre 2003.

³Un job est un ensemble de processus qui participent à l'accomplissement d'une tâche.

⁴telle que la moyenne, l'écart-type ou la variance de quelques variables