

# 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 CGG<sup>1</sup> 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'architecture<sup>2</sup>. 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 jobs<sup>3</sup>, 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 base<sup>4</sup> souvent rencontrée dans ce type de cas.

---

<sup>1</sup>Compagnie Générale de Géophysique

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

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

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

## 1 Mise en forme et caractéristiques du problème

La caractérisation de jobs vise à fournir aux administrateurs d'un centre de calcul une information claire sur les types de jobs exécutés. L'objectif est donc de déterminer les propriétés les plus marquées des jobs. Ainsi certains peuvent nécessiter une grande puissance de calculs, d'autres effectuent beaucoup d'accès aux données ou exploitent largement le réseau, à différentes étapes du processus. Nous voulons donc mieux connaître les jobs, dans ce qu'ils ont de plus caractéristique, les classer.

Pour cela, nous disposons d'une plateforme d'acquisition. Elle est principalement composée d'une base de données alimentée par un collecteur, dont la tâche est de récupérer les mesures issues des machines composant le parc de production. Les statistiques sont collectées sur chaque machine par un capteur qui communique périodiquement au collecteur les informations mises à jour. Différents types d'informations peuvent être remontés par des capteurs spécifiques : les *variables d'exécution* (temps CPU utilisé, mémoire utilisée, nombre de Ko échangés via le réseau ou une unité de stockage, ...).

L'exploitation de ces données peut s'effectuer à l'aide de techniques se séparant naturellement en deux. Chacune répond à un besoin particulier dans la mise en évidence de la charge de travail [Elnaffar et Martin, 2002, Calzarossa et Serazzi, 1993] :

- les **techniques statiques** sont employées afin d'explorer les caractéristiques intrinsèques des charges de travail (moyennes, écarts-types, corrélations, ...).
- les **techniques dynamiques** tentent de décrire le comportement de la charge, c'est à dire son évolution au cours du temps (approximations par l'historique, distributions statistiques, réseaux de neurones, chaînes de Markov, ...).

Ces deux types de techniques sont complémentaires, puisque la plupart du temps c'est leur conjonction qui permet d'obtenir le résultat souhaité concernant la caractérisation.

## 2 Architecture de la solution d'extraction

### 2.1 Modèles de Markov et SVM

Les modèles de Markov sont des graphes dont chaque noeud représente un état et chaque arc orienté la probabilité de passer d'un état à un autre. Ils sont utilisés ici afin de décrire la dynamique d'un job, les états étant les phases dans lesquelles se retrouve un job ( $E = \{\text{utilisation intensive du CPU, de la mémoire, du réseau}\}$ , par exemple). On pourra se reporter à [Elnaffar et Martin, 2002] pour plus d'informations. Enfin, ils nous permettent de décrire un job sur une mainframe ou un cluster (en utilisant un méta-modèle).

Les SVM<sup>5</sup> possèdent des qualités de classification et régression qui les placent devant d'autres méthodes, si elles sont bien utilisées [Cornuéjols, 2002]. Leur champ d'application, pour la problématique nous intéressant, est la détection de phases. Sachant que les jobs possèdent des phases assez marquées, les SVMs permettent de bien les distinguer. Nous nous intéressons à des paquets de données dans une fenêtre  $[t + n\tau, t + (n + 1)\tau]$  avec  $n \in \mathbb{N}$ . L'objectif est d'attribuer à chaque paquet de données

---

<sup>5</sup>Support Vector Machines ou machines à supports de vecteurs

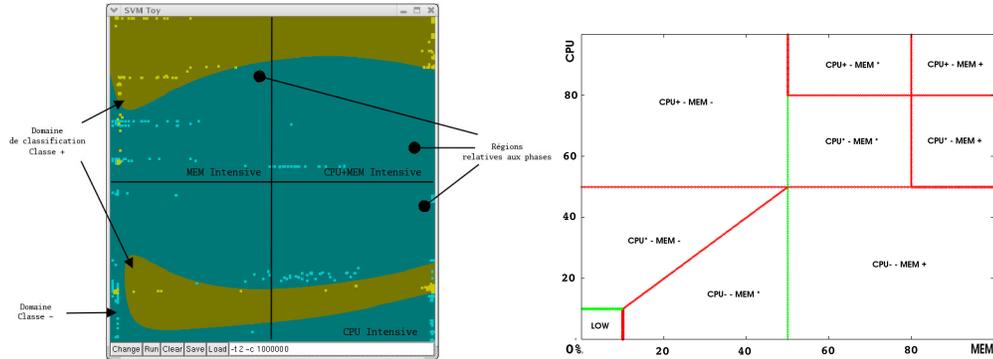


FIG. 1 – Régions de l’espace pour la détermination de phases (4 simplistes) et Domaines CPU, MEM avancés

une phase parmi l’ensemble de celles définies dans un ensemble  $E$  d’états (= phases). Pour cela, on s’intéresse à la manière dont une SVM peut séparer les données.

### 2.1.1 Déroulement du processus de caractérisation avec SVM

Tout d’abord, les phases [Cook *et al.*, 2001] nous intéressant doivent être construites, ce qui revient à décrire des domaines dans un espace dont les dimensions sont les variables considérées. Des exemples sont donnés Figure 1. Puis de manière chronologique les actions suivantes sont effectuées pour un paquet  $n$  :

1. suivant, la taille du paquet,  $\tau$  :
  - (a) si la taille est supérieure à un minimum, on détermine le domaine (ou zone de prédominance) du paquet  $n$ , en appliquant une SVM avec comme classe– une répartition uniforme de points et comme classe+ les mesures contenues dans le paquet  $n$ . On obtient alors le *domaine de ce paquet*.
  - (b) si la taille du paquet est trop faible, on sélectionne la phase à laquelle appartient le plus grand nombre de points de mesure et on l’attribue au paquet (détection naïve).
2. si le paquet  $n$  est différent du paquet  $n-1$  de part la moyenne et l’écart-type des mesures qu’il contient, alors on applique une SVM avec comme classe– les mesures du paquet  $n-1$  et comme classe+ les mesures du paquet  $n$ . Le domaine ainsi obtenu, marque la *zone de prédominance relativement au paquet précédent*. On ajoute un pourcentage du domaine ainsi obtenu à celui trouvé en 1.
3. Ensuite 2 cas apparaissent :
  - (a) si le domaine<sup>6</sup> est contenu, au delà d’un taux limite, dans une phase définie, alors le paquet se voit attribuer cette phase.

<sup>6</sup>Synthèse du domaine obtenu en 1 et celui obtenu par opposition au paquet  $n-1$ .

- (b) si le domaine ne répond pas au cas précédent, on coupe le paquet en 2,  $\tau' = \frac{\tau}{2}$  et on retourne en 1.

Ceci doit permettre de détecter les phases les plus marquées, tout en prenant en compte les phases de transition, mais sans détecter des changements de phase lorsque quelques mesures ne sont pas homogènes avec le reste d'un paquet. *Ces contraintes sont respectées grâce à l'utilisation de paquets de longueur variable et l'application d'une SVM entre 2 paquets statistiquement différents.*

## 2.2 Réalisation du modèle de Markov enrichi

Une fois les domaines, décrivant les phases, définis à l'aide des variables sélectionnées, nous appliquons la détection, à l'aide de la SVM, (cf 2.1.1) sur les mesures. A partir de cela nous pouvons calculer les probabilités de transition entre les phases et construire le modèle de Markov. Pour chacune des phases retenues, les informations statiques collectées (durée d'une phase, nombre de paquets relatifs à cette phase, moyennes et écart-types des variables considérées, nouvelles variables issues de l'ACP, caractéristiques d'une clusterisation, ...) seront regroupées afin de donner un ordre d'importance de chaque phase au regard de l'exécution du processus. *C'est cette association de résumés ciblés<sup>7</sup> de données à chaque phase qui constitue le modèle de Markov enrichi.*

Le résultat est donc une synthèse des *variables d'exécution* du job qui tient compte de sa dynamique, et suivant laquelle des informations peuvent être détaillées suivant différents points de vue.

## 2.3 Classification supervisée

Le but est de classifier les jobs grâce aux modèles de Markov enrichis qui nous donnent leurs caractéristiques, si possible de manière indépendante du système hôte, via les *variables structurelles*<sup>8</sup>. Pour cela, nous proposons d'utiliser les réseaux bayésiens naïfs [Rish, 2001], qui devront permettre de répartir les jobs suivant des classes définies par l'utilisateur. Les faits étant du type job intensif en CPU, en mémoire, ... (relatifs à l'exécution) et machine à faible ou fort rendement en CPU, accès disque, ... (relatifs à la structure). Enfin, l'objectif second est de faciliter la mise en place d'une classification basée sur le raisonnement à partir de cas.

## 3 Résultats

La SVM utilisée est une C-SVC qui s'applique bien au cas binaire qui nous intéresse. Le noyau est du type RBF<sup>9</sup> dont le paramètre vaut :  $\gamma$  = taille des échantillons. Le coût des exemples mal classés :  $c = 10000$ , cette valeur permet d'obtenir des zones de prédominance de taille correcte, tout en prenant en compte les cas où les mesures ne sont pas très homogènes [Cornuéjols, 2002, Wahba *et al.*, 2001]. Le seuil d'attribution

<sup>7</sup>suivant certaines variables

<sup>8</sup>indépendantes du temps, elles correspondent aux caractéristiques du matériel (type et fréquence du CPU, taille mémoire, ...)

<sup>9</sup>Radial Basis Function (de la forme  $\exp(-\gamma * |u - v|^2)$ )

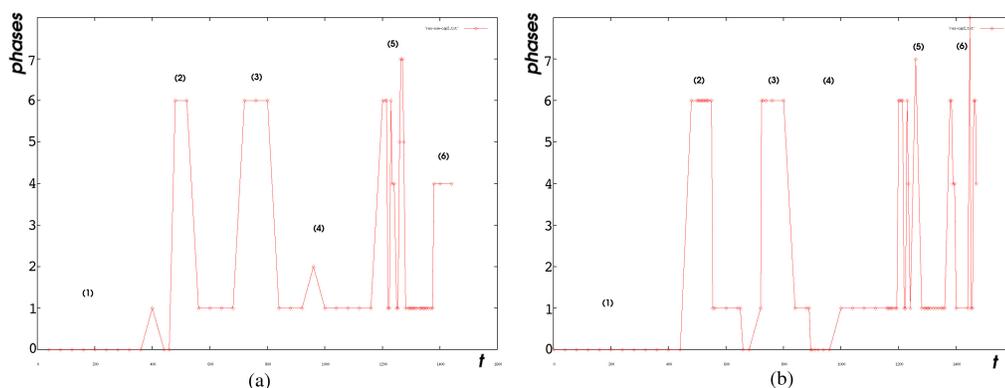


FIG. 2 – Phases avec (a) et sans (b) SVM

d'une phase à un paquet a été fixé à 50%. De ce fait, si la moitié des entités considérées se trouve dans un domaine alors la phase correspondante est attribuée au paquet.

Les données se composent de 1476 mesures recueillies sur une durée d'environ 16h30, relatives à un job tournant sur une SGI Origin2000. Cette machine possède 8 processeurs MIPS mais l'étude est restreinte à un seul de ceux utilisés.

Nous comparons, ici, 2 types de détecteurs :

- détecteur naïf : Pour un paquet relativement homogène, il comptabilise le nombre de mesures appartenant à chaque domaine. Le domaine contenant le plus grand nombre de mesures est sélectionné, et le paquet se voit affecté de la phase associée. Si l'écart-type des mesures d'un paquet est trop important, celui-ci est scindé en 2 jusqu'à ce qu'il soit plus compact.
- détecteur SVM : cf 2.1.1.

Les phases utilisées sont définies sur la Figure 1 avec 3 grands degrés d'importance (par ordre décroissant) : +, \*, -. Les correspondances entre la valeur en ordonnée (Figure 2) et la phase sont les suivantes : 0 : LOW, 1 : CPU- - MEM\*, 2 : CPU\* - MEM -, 3 : CPU\* - MEM+, 4 : CPU+ - MEM\*, 5 : CPU- - MEM+, 6 : CPU+ - MEM-, 7 : CPU\* - MEM\*, 8 : CPU+ - MEM+ .

Pour la série de mesures "stables", les phases en (1), (2) et (3) sont facilement et correctement détectées par les détecteurs, cependant le détecteur sans SVM annonce une phase d'activité très faible (entre (2) et (3) sur la Figure 2) qui n'apparaît pas avec l'autre. C'est un cas limite où les mesures sont très proches de la frontière du domaine (LOW) et globalement, l'activité se situe en dehors de cette phase. De plus le saut vers la phase 2 en (4) observé avec la SVM est justifié, du fait d'une activité CPU oscillante sur toute la durée du paquet. Ce résultat est dû à la séparation avec le paquet précédent effectuée grâce à la SVM, qui favorise le domaine ainsi choisi.

Pour la série de mesures "variables", entre (5) et (6), le détecteur sans SVM annonce plusieurs phases en un court laps de temps. Même si on retrouve effectivement des mesures correspondant, ce n'est pas la tendance moyenne, celle donnée par le détecteur avec SVM. Les phases y sont plus "stables", plus caractéristiques de l'activité.

Enfin, en effectuant un tri sur les phases dans lesquelles le job passe le plus de temps

et dont les probabilités de transition sont les plus élevées, nous mettons en évidence, d'un point de vue dynamique, une phase centrale : "CPU- - MEM\*" et une secondaire dans laquelle se retrouve assez souvent le job : "CPU+ - MEM-".

## Conclusion

La caractérisation de jobs met en jeu la productivité des machines, une optimisation des ressources; c'est une approche plus raisonnée de la puissance de calcul. L'utilisation de différentes techniques d'analyse est requise pour une caractérisation complète. Ainsi, les modèles de Markov enrichis sont présentés ici comme la base de toute caractérisation globale des jobs. En effet, ils concentrent les aspects statiques et dynamiques de l'exécution au sein d'une même représentation. De part leur intelligibilité aisée, ils sont utilisés à l'échelle d'une ressource ou d'un cluster (méta-modèle). A partir de cela, l'utilisation d'un réseau bayésien tenant compte de ces données permettrait une caractérisation absolue<sup>10</sup>. Le prototype développé a montré que la détection de phases à l'aide d'une SVM engendre des modèles de Markov de qualité, reflétant le déroulement du job.

## Références

- [Calzarossa et Serazzi, 1993] M. Calzarossa et G. Serazzi. Workload characterization : A survey. In *Proceedings of the IEEE*, volume 81, pages 1136–1150, August 1993.
- [Cook et al., 2001] J. Cook, E.E. Johnson, et R.L. Oliver. Examining performance differences in workload execution phases. In *Proceedings of the 4th IEEE International Workshop on Workload Characterization*, pages 82–90, December 2001.
- [Cornuéjols, 2002] A. Cornuéjols. Une nouvelle méthode d'apprentissage : Les svm. *Bulletin de l'AFIA*, 2002.
- [Elnaffar et Martin, 2002] S. Elnaffar et P. Martin. Characterizing computer systems' workloads. *ACM Computing Surveys Journal*, December 2002.
- [Rish, 2001] I. Rish. An empirical study of the naive bayes classifier. In *Workshop on 'Empirical Methods in AI' IJCAI*, 2001.
- [Wahba et al., 2001] G. Wahba, Y. Lin, Y. Lee, et H. Zhang. Optimal properties and adaptive tuning of standard and nonstandard support vector machines. Technical Report No. 1045, October 2001.

## Summary

The *workload characterization of jobs execution* is archived through *static* and *dynamic* analysis of various measures collected on each computer. SVM are used for detecting phases and naive Bayesian networks automatize the analysis of enhanced Markov models. This ones are introduced as a base for formal and synthetic descriptions of jobs on batch/parallel systems. Finally, some results are discussed.

---

<sup>10</sup>indépendante de l'environnement d'exécution