

Predictibilité et efficacité des traitements concurrents dans les systèmes multicœurs

Unité d'accueil : Institut de Recherche en Informatique de Toulouse

Equipe : Acadie

Encadrement : Philippe Quéinnec Tel : 0534322184

Encadrement : Philippe Mauran Tel : 0534322183

Contexte

Afin de répondre à leurs besoins toujours croissants des applications en termes de puissance de calcul, les applications temps réel peuvent espérer profiter du potentiel des architectures multi-cœurs sur cet aspect. Cependant, l'évaluation fine de contraintes temporelles strictes pour l'exécution de traitements concurrents est difficile dans ce cadre, du fait de l'indéterminisme résultant des conflits d'accès aux ressources partagées : ainsi, lorsque les conflits sont résolus au moyen de primitives de synchronisation, le pire temps d'exécution correspond souvent à une exécution séquentielle, mono-cœur. Par ailleurs, les solutions basées sur la synchronisation présentent l'inconvénient de rester largement artisanales en termes de conception (avec les inconvénients qui en résultent : risque d'erreurs, nécessité d'une expertise...) de ne pas être efficaces à gros grain (puisqu'elle supprime aussi le parallélisme, pour la durée du traitement des conflits), et de ne pas être compositionnelles/modulaires à grain fin (chaque évolution/combinaison avec d'autres activités remet en cause la solution précédente)

Assez naturellement, des voies alternatives ont été explorées, qui visent à permettre l'occurrence de conflits, dans la mesure où leur effet est jugé tolérable (protocoles de cohérence mémoire), ou peut être corrigé (protocoles transactionnels. Cette approche fait d'une part référence au résultat, c'est-à-dire à l'état de la ressource/donnée partagée, et d'autre part à des critères (de cohérence) autorisant à permettre un conflit. Elle est réalisée à partir du contrôle des accès à l'objet partagé, à grain fin (protocoles de cohérence de données dupliquées) ou à gros grain (mécanismes transactionnels). Cette approche présente l'intérêt d'autoriser potentiellement davantage de parallélisme (moins de blocages), et d'être compositionnelle (les critères de cohérence ne dépendent pas des applications). Elle a l'inconvénient d'être plus coûteuse dans sa mise en œuvre, et de ne pas toujours être possible dans le cas général, selon les choix de mise en œuvre (ainsi, un système en interaction avec un environnement qu'il ne contrôle pas, ne pourra pas forcément corriger une action passée)

Thème de travail

La thématique de travail proposée s'inscrit dans cette dernière direction, en considérant les protocoles optimistes de mémoire transactionnelle, qui ouvrent donc la perspective d'une amélioration du degré de concurrence, de la facilité de conception, d'adaptation et d'évolution des applications, mais peuvent dégrader encore davantage l'évaluation de contraintes temporelles, puisque, dans le cas général, l'occurrence de conflits peut, dans le pire cas, amener à annuler et relancer une transaction un nombre de fois indéterminé. Nous nous intéressons donc à l'évaluation et l'amélioration de la prédictibilité (en termes de temps ou de vivacité) de l'exécution d'un ensemble de traitements concurrents en interaction via une mémoire partagée selon des protocoles sans blocage, comme la mémoire transactionnelle.

Objet de la thèse

Le point de départ du projet sera l'analyse et l'évaluation des protocoles existants du point de vue de l'efficacité et de la prédictibilité. Les recherches actuelles mettent en évidence la variété des notions de progrès et d'équité, en rapport essentiellement avec la prise en compte des défaillances des processus. Dans l'optique d'employer ces protocoles pour des systèmes temps réel embarqués, il faudra étudier les possibilités de

- coupler le modèle d'exécution de ces protocoles (basé sur un un temps logique) avec le temps physique, afin de permettre d'évaluer la satisfaction de propriétés temps réel (comme la certitude ou la probabilité d'échéance à une date donnée) sur les activités s'exécutant en concurrence sur un multicœur.
- tirer parti des spécificités du contexte temps réel embarqué (tâches statiques, régulières, possibilités de synchronisme partiel) pour réduire effectivement l'incertitude quant aux propriétés temporelles de l'exécution.

À partir de cette analyse, il s'agira

- d'adapter et étendre ces protocoles de gestion de la concurrence sans blocage pour assurer ou renforcer les propriétés d'efficacité et de prédictibilité, Notre approche est de jouer sur l'ordonnement pour contrôler et limiter les conflits. A cette fin, deux pistes complémentaires sont envisagées :
 - prédire les conflits, à partir d'une analyse des transactions, ou d'heuristiques sur leur comportement passé. Le résultat de cette prédiction est abstrait par un graphe de conflits. En l'état actuel de nos travaux, ce graphe est juste défini a priori, et sert de base pour le point suivant. Il serait intéressant de définir et d'intégrer des heuristiques de prédiction des conflits pour guider l'élaboration du graphe.
 - utiliser la connaissance sur les conflits pour évaluer et garantir des contraintes temporelles ou de vivacité sur les traitements. Notre principale contribution sur ce plan est de définir un algorithme d'ordonnement garantissant la vivacité pour les traitements, ainsi que des bornes pour le temps d'exécution des traitements dans ces conditions. Cet algorithme doit pouvoir être amélioré, pour mieux tirer parti de la structure des conflits afin de réduire les temps d'exécution (en moyenne et dans le pire cas).
- et/ou d'explorer un cadre de programmation concurrente prédictible, visant à
 - faciliter la conception de programmes concurrents en environnement multicœur, en définissant des structures de contrôle/d'organisation pour les transactions, et en permettant la composition entre protocoles de cohérence, et entre transactions et synchronisation classique.
 - améliorer l'efficacité de l'exécution transactionnelle, par la mise en œuvre de protocoles adaptatifs/hybrides,
 - expliciter la relation entre tolérance aux fautes et contrôle de concurrence, en particulier de sorte à offrir au programmeur des garanties de terminaison (de manière prédictible en temps-réel ou non), ce qui semble raisonnable pour un ensemble de traitements de nature systolique, caractéristiques des systèmes temps-réel embarqués.

Cadre méthodologique

Les études menées s'appuient avant tout sur des modèles formels comme les systèmes de transition et la logique temporelle. Notre cadre de prédilection actuel est le formalisme et l'environnement TLA, proposés par Leslie Lamport. Les preuves développées peuvent être outillées dans ce cadre, ou en s'appuyant sur les compétences de l'équipe dans le domaine des assistants de preuve comme Coq.

Par ailleurs, il est également possible, à des fins d'illustration, d'exploration préliminaire, ou d'expérimentation, de recourir à une plate-forme de simulation qui a été développée pour l'occasion, et qui peut permettre d'apprécier la finesse des prédictions, et l'influence des paramètres temporels, sur l'exécution des traitements concurrents.

Enfin, les développements applicatifs éventuels s'appuieront sur la plateforme Java.