Predição de Desempenho de Aplicações Paralelas para Máquinas Agregadas Utilizando Modelos Estocásticos

Lucas Janssen Baldo, Luiz Gustavo Leão Fernandes

Pontifícia Universidade Católica do Rio Grande do Sul
Av. Ipiranga, 6681 prédio 16 - PPGCC
{lbaldo, gustavo}@inf.pucrs.br

Introdução

Um dos maiores problemas na área de computação de alto desempenho é a dificuldade de definir qual a melhor estratégia de paralelização de uma aplicação. Neste contexto, a utilização de métodos analíticos para a avaliação de desempenho de aplicações paralelas aparece como uma alternativa interessante para auxiliar no processo de escolha das melhores estratégias de implementação. Neste trabalho, propõe-se a adoção do formalismo de Redes de Autômatos Estocásticos para modelar e avaliar o desempenho de aplicações paralelas especialmente desenvolvidas para máquinas agregadas (i.e., *clusters*). Trabalhos anteriores similares são muito genéricos [Meira 95], ou específicos demais [Gemund 03]. A metodologia utilizada é baseada na construção de modelos genéricos para descrever esquemas clássicos de implementação paralela, tais como Mestre/Escravo, Fases Paralelas, *Pipeline* e Divisão e Conquista. Estes modelos são adaptados em casos de aplicações reais através da definição de valores para parâmetros de entrada dos modelos. Finalmente, com intuito de verificar a precisão da técnica de modelagem adotada, comparações com resultados de implementações reais são apresentadas.

Modelos SAN Genéricos

Nesta seção, os modelos SAN genéricos para agregados de máquinas são apresentados, focando no processo de troca de dados e tempo de processamento. A relação entre essas duas características são o ponto chave para determinar o sucesso de implementações paralelas.

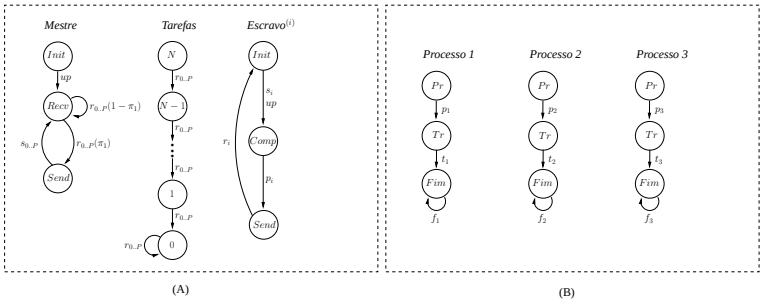

Figura 1: Modelos SAN genéricos.

O modelo SAN para mestre/escravo é mostrado na Figura 1 (A). Este modelo contém um autômato *Mestre*, um autômato *Tarefas*, e P autômatos *Escravos*$^{(i)}$ ($i = 1..P$). O modelo de fases paralelas é apresentado na Figura 1 (B). Este modelo possui P autômatos *Processos*$^{(i)}$ ($i = 1..P$). Os resultados obtidos através da resolução destes dois modelos são comparados com os tempos de execução obtidos através das implementações reais das duas aplicações (Figura 2).

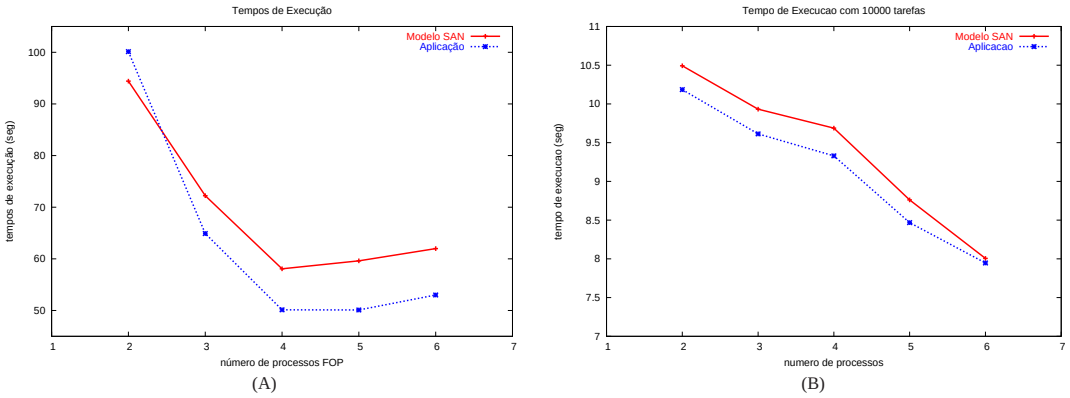

Figura 2: Alguns resultados obtidos com os modelos SAN genéricos.

Pode-se notar que os modelos SAN conseguem detectar características comportamentais semelhantes às obtidas pela aplicação como por exemplo o ponto de inflexão apresentado na Figura 2 (A).

Conclusão

O uso de um formalismo estocástico para modelar sistemas complexos parece fornecer vantagens em comparação a outras modelagens matemáticas. Nós acreditamos que modelos SAN são uma boa abordagem para contornar este problema e o estudo de modelagem apresentado neste trabalho para modelar aplicações para ambientes de alto desempenho, ou mais especificamente máquinas agregadas, pôde ser desenvolvido por não especialistas em avaliação de desempenho. Este fato reforça a facilidade do uso de modelos SAN para prever desempenho de aplicações em diferentes áreas. Dois trabalhos sobre este assunto foram submetidos e aceitos na comunidade científica, dentre eles destaca-se [BBFFS 05].

Referências

[BBFFS 05] Baldo, L; Brenner, L; Fernandes, L G; Fernandes, P; Sales, S. **Performance Models for Master/Slave Parallel Programs.** Electronic Notes In Theoretical Computer Science, v.128, n.4, p.101–121, Abril 2005.

[Meira 95] W. Meira Jr. **Modeling Performance of Parallel Programs.** Techinical Report, n. 589, University of Rochester, 1995.

[Gemund 03] A. J. C. van Gemund **Symbolic Performance Modeling of Parallel Systems.** IEEE Transactions on Parallel and Distributed Systems, v.14, n.2, p.154-165, February 2003.