

Escalonamento e simulação de ambientes de Computação Voluntária

Bruno Donassolo¹, Cláudio Geyer¹

¹Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brasil

{blmdonassolo,geyer}@inf.ufrgs.br

1. Introdução

A computação voluntária é uma plataforma global cujo uso vem aumentando consideravelmente nos últimos anos devido, principalmente, ao desenvolvimento das redes de computadores e à diminuição da relação custo/desempenho dos componentes computacionais. Seus recursos são formados por máquinas disponibilizadas por pessoas voluntárias, isto é, que doam uma parte de seu poder computacional espontaneamente. Consequentemente, é caracterizada pela grande volatilidade e heterogeneidade de seus componentes [Kondo et al. 2007].

Entre as diversas plataformas desse tipo foram desenvolvidas nos últimos anos, foi o BOINC [Anderson 2004] que alcançou a maior representatividade, sendo largamente utilizado por diversos projetos de pesquisas científicas. Nessa área, um dos grandes desafios é como escalonar as tarefas entre os clientes de forma a obter o melhor desempenho e a finalização das tarefas dentro do seu limite de tempo. Porém, o estudo dos algoritmos de escalonamento numa plataforma real é dificilmente praticável devido aos custos implicados. Dessa forma, o uso de simuladores é imprescindível, permitindo o estudo e a verificação do seu desempenho em plataformas complexas antes da sua real implantação.

2. Trabalhos Relacionados

Diversos trabalhos vem sendo desenvolvidos ultimamente no estudo das políticas de escalonamento dos clientes e servidores do BOINC. Entre eles, podemos citar o trabalho de Estrada et al. [2006] que propõe um escalonamento considerando uma classificação prévia dos clientes. Ainda, o artigo de Anderson e Reed [2009] mostra as técnicas usadas atualmente para tratar a diversidade das tarefas e máquinas nesses ambientes.

Esses trabalhos normalmente utilizam simuladores, a maior parte construídos somente para satisfazer as necessidades do estudo proposto. Entre eles, o Simba [Taufer et al. 2007], EmBOINC [Estrada et al. 2009], cujo objetivo é estudar as políticas de escalonamento do lado servidor do BOINC, e o BOINC Client Simulator, o qual é usado para estudar os algoritmos de escalonamento locais ao cliente.

Entretanto, esses simuladores possuem restrições importantes, como a simulação de apenas um cliente/servidor ou a dificuldade do teste de novas políticas de escalonamento. Como alternativa para transpor algumas das restrições acima citadas, existem simuladores de propósito geral, no qual o usuário descreve completamente o comportamento da sua aplicação. Um exemplo nessa categoria é o simulador SimGrid, cujo objetivo é o estudo de aplicações distribuídas e que, apesar da capacidade de simular até 200.000 nós numa máquina, possui problemas na velocidade da simulação, conforme descrito no trabalho de Hein et al. [2008].

3. Objetivos

O primeiro objetivo desse trabalho é realizar as modificações necessárias no simulador SimGrid para que este possua um desempenho razoável na simulação de aplicações de computação voluntária e portanto, viabilizar os estudos desejados no segundo objetivo. Em detalhes, essas mudanças incluem a modificação no modelo de simulação da CPU que se mostrou ineficaz quando confrontado aos tipos de aplicações típicos desta área.

O segundo objetivo é estudar o impacto da interação de projetos com diferentes tipos de tarefas e diferentes esquemas de escalonamento. Mais precisamente, verificar o impacto de projetos com tarefas em rajadas, curtas e raras, no desempenho de projetos com aplicações longas e constantes. Ainda, estudar como o uso de muitas réplicas, com o intuito de diminuir o tempo de resposta das tarefas dos projetos curtos, prejudica o número total e o tempo de conclusão das tarefas dos diferentes projetos.

4. Trabalho em Andamento

O trabalho em andamento atualmente é a adaptação do simulador SimGrid para a simulação de ambientes de computação voluntária. Este trabalho está sendo realizado durante um estágio junto a um dos desenvolvedores do SimGrid no laboratório LIG em Grenoble. Os resultados preliminares mostram um ganho de desempenho do simulador numa situação típica (vários clientes executando tarefas de um servidor) de algumas ordens de grandeza. Além disso, está sendo realizado a modelização do comportamento dos clientes e servidores BOINC utilizando uma das interfaces disponíveis, chamada MSG. Esses passos permitirão então, realizar os estudos comparativos descritos nesse artigo.

Referências

- Anderson, D. P. (2004). Boinc: a system for public-resource computing and storage. In *GRID '04*, pages 4–10.
- Anderson, D. P. and Reed, K. (2009). Celebrating diversity in volunteer computing. In *HICSS*, pages 1–8.
- Estrada, T., Flores, D. A., Taufer, M., Teller, P. J., Kerstens, A., and Anderson, D. P. (2006). The Effectiveness of Threshold-Based Scheduling Policies in BOINC Projects. In *E-SCIENCE '06*, page 88.
- Estrada, T., Taufer, M., Reed, K., and Anderson, D. P. (2009). Emboinc: An emulator for performance analysis of boinc projects. In *IPDPS*, pages 1–8.
- Heien, E. M., Fujimoto, N., and Hagihara, K. (2008). Computing low latency batches with unreliable workers in volunteer computing environments. In *IPDPS*, pages 1–8.
- Kondo, D., Fedak, G., Cappello, F., Chien, A. A., and Casanova, H. (2007). Characterizing resource availability in enterprise desktop grids. *Future Gener. Comput. Syst.*, 23(7):888–903.
- Taufer, M., Kerstens, A., Estrada, T., Flores, D., and Teller, P. J. (2007). SimBA: A Discrete Event Simulator for Performance Prediction of Volunteer Computing Projects. In *PADS '07*, pages 189–197.