



Marque a opção do tipo de trabalho que está inscrevendo:

Resumo

Relato de Caso

FRAMEWORK PARA EXECUÇÃO PARALELA DE MODELOS DE SIMULAÇÃO DE CULTURAS.

AUTOR PRINCIPAL: Angela Mazzonetto.

CO-AUTORES: Marcelo Trindade Rebonatto.

ORIENTADOR: Carlos Amaral Hölbig.

UNIVERSIDADE: Universidade de Passo Fundo.

INTRODUÇÃO

Dados climáticos e meteorológicos estão disponíveis nas mais diversas bases e uma das áreas de maior aplicação é a área agrícola, onde modelos de simulação de culturas e doenças fazem uso destes dados. Quando o modelo necessita de uma grande quantidade de dados observados devido a sua área de abrangência ou ao período contemplado por estes dados com objetivo de abranger uma maior região isso torna a execução dos modelos repetitiva, pois o mesmo modelo tende a ser executado muitas vezes. Devido a este fato, é necessária a escolha de alternativas computacionais que proporcionem um ganho no tempo de execução destes modelos, principalmente quando manipulam uma grande quantidade de dados observados. Entre as alternativas avaliadas por esta pesquisa destaca-se neste trabalho o uso do processamento paralelo, e que poderia ser utilizado para a execução dos modelos.

DESENVOLVIMENTO:

O modelo de simulação escolhido para ser estudo de caso deste trabalho é o modelo de simulação de culturas e doenças que é utilizado no grupo de pesquisa Mosaico da Universidade de Passo Fundo e que necessita de uma melhora no seu tempo de execução pois esta é uma das limitações atuais estão impossibilitando sua maior aplicação. Este modelo, nomeado Cropsim-Wheat, foi desenvolvido por Leslie A. Hunt [1] e atualmente pertence ao sistema Decision Support System for Agrotechnology Transfer (DSSAT) que é composto por diversos modelos de simulação. O Cropsim simula o crescimento e desenvolvimento da cultura do trigo, com dados de entrada que correspondem às características da planta e dados meteorológicos da localidade.

A paralelização deste modelo se deu com uma abordagem da metodologia de mestre-escravo que, de acordo com a literatura, é a mais comumente utilizada nas aplicações paralelas

em um ambiente computacional de um cluster de computadores. A abordagem de paralelização da execução funciona de forma que um conjunto de comandos de execução do modelo seja enviado para cada processador. Cada linha de comando contém a execução de um tratamento, o qual tratamento corresponde a um conjunto de dados referentes à localidade. Portanto, por exemplo, se são necessárias 9600 execuções do modelo (correspondendo a quatro localidades com 2400 tratamentos cada uma), serão enviadas 100 linhas de comandos de execução para cada processo. Este envio é feito pelo processador mestre. Logo após, o processador escravo começa as execuções e, quando este chega ao final das 100 linhas de comando recebidas, envia um aviso ao mestre que finalizou suas execuções. Se ainda houver linhas de comando a serem executadas, o mestre as envia aos seus escravos até que não as tenha mais. Quando todo o trabalho estiver finalizado, o mestre envia uma mensagem de fim de trabalho aos escravos. Esta troca de mensagens é feita pelo padrão Message Passing Interface (MPI), que é uma interface de programação para comunicação de dados entre processos paralelos. Esta ferramenta oferece infraestrutura para modalidades de computação paralela quando existe a necessidade de passar informações entre os vários processadores ou nodos de um cluster com memória distribuída [2].

Para que fosse possível obter ganhos mais significativos, o modelo cropsim foi executado, também, utilizando memória distribuída, ou seja, cada lote de simulações teve seus arquivos de dados de entrada armazenados no computador em que fosse realizar a execução do processamento dos dados. Foi realizada a simulação com memória distribuída pelo fato de que o cropsim utiliza muita escrita e leitura de dados, o que pode acarretar em concorrência na rede pela troca de pacotes contendo dados.

CONSIDERAÇÕES FINAIS:

Após execução dos testes no cluster com memória distribuída, utilizando até 64 processadores, pôde-se observar ganho no desempenho e concluir que a paralelização da execução de modelos de simulação em um cluster de computadores é vantajosa, pois permite ampliar a quantidade de dados e de localidades a serem processados pelo modelo de simulação.

REFERÊNCIAS

[1] HUNT, L. A.; PARARAJASINGHAM, S. Cropsim-wheat: A model describing the growth and development of wheat. Canadian Journal Plant Science, Department of Crop Science, University of Guelph, Guelph, Ontario, Canada, 1995. Disponível em: <<http://pubs.aic.ca/doi/pdf/10.4141/cjps95-107>>.

[2] LI, T. et al. Dynamic parallelization of hydrological model simulations. Environmental Modelling and Software, Elsevier Ltd, v. 26, n. 12, p. 1736-1746, 2011. ISSN 13648152. Disponível em: <<http://dx.doi.org/10.1016/j.envsoft.2011.07.015>>.