Avaliação do Tempo de Processamento e Comunicação via Rotinas MPI Assíncronas no Modelo OLAM *

Matheus Beniz Bieger[1], **Claudio Schepke**[1]

[1]Curso de Ciência da Computação
Universidade Federal do Pampa – Campus Alegrete
Av. Tiarajú, 810 – Bairro Ibirapuitã – Alegrete – RS – CEP: 97546-550

matheusbieger@alunos.unipampa.edu.br, claudioschepke@unipampa.edu.br

Resumo. *A previsão do tempo é um fator importante na sociedade. Para estas previsões, são utilizados modelos de previsão numérica que simulam as condições atmosféricas atuais. Neste contexto, esse artigo avalia o tempo de comunicação e computação do modelo OLAM utilizando a interface de programação MPI. Além disso, realiza um estudo sobre as rotinas de comunicação assíncronas utilizadas no modelo.*

1. Introdução

Os modelos numéricos de previsão do tempo e clima estão sendo largamente utilizados nas últimas décadas. Isso se deve ao fato de haver discussões recorrentes sobre as mudanças climáticas e a previsão de cenários futuros. Modelos de previsão numérica auxiliam na detecção e prevenção de catástrofes, aquecimento global, temperatura, entre outros fatores relevantes para a comunidade global [Silva et al. 2009].

O modelo OLAM (Ocean-Land-Atmospheric Model) [Walko and Avissar 2008] é um modelo de previsão numérica que permite a realização de simulações no âmbito global e regional, utilizando um refinamento de malha dinâmico. Este modelo possui um tempo de execução significativo. Assim sendo, busca-se acelerar a sua execução utilizando técnicas de programação paralela. Visando este aumento de desempenho, os desenvolvedores do modelo implementaram uma versão paralela do código, utilizando a interface de programação MPI. Para maximizar o desempenho é aconselhado otimizar ao máximo a utilização do MPI. Em [Pydd et al. 2014] já foram realizados testes quanto ao desempenho do MPI nessa aplicação, contudo não foi realizada uma análise mais detalhada das rotinas MPI utilizadas.

Este trabalho realiza uma análise das chamadas de comunicação MPI que ocorrem entre as iterações do modelo, a fim de estudar e avaliar o custo computacional dessas chamadas. O estudo do perfil de execução das chamadas MPI mostra que na inicialização do modelo tais chamadas são irrelevantes e na execução, podem ocorrer variações no tempo de comunicação decorrentes da chamada de rotinas assíncronas.

Este artigo está organizado em 5 seções. A Seção 2 aborda os trabalhos relacionados a este estudo. A Seção 3 apresenta o modelo OLAM e suas particularidades. Os testes estão descritos na Seção 4. A Seção 5 analisa e realiza inferências sobre os resultados obtidos. Por fim, a Seção 6 apresenta as conclusões e direções futuras.

*Este trabalho foi fomentado através do Edital No 05/2014 PIBIC - CNPq / UNIPAMPA.

2. Trabalhos Relacionados

A aplicação de paralelismo na execução de modelos atmosféricos é amplamente estudada na comunidade científica. Há uma diversidade de modelos existentes e as mais variadas técnicas de programação paralela são aplicadas.

Em [Schepke et al. 2010] é realizada uma avaliação de performance do modelo atmosférico OLAM em um ambiente multicore. Neste trabalho, os autores avaliaram a escalabilidade estudando o *speedup* obtido na utilização de diferentes configurações de nós, processadores e núcleos de um cluster. Os resultados indicaram que na utilização de muitos *cores* do cluster, há uma maior competição para o acesso a memória e mais *caches misses*. Em [Pydd et al. 2014] os autores avaliaram o desempenho paralelo do modelo OLAM utilizando processos MPI em memória compartilhada. Foram realizados testes com resoluções do modelo e número de processos distintos. As execuções foram feitas utilizando 1, 2, 4, 8, 16 e 32 processos. Os resultados apontaram que o tempo de inicialização e refinamento do modelo são significantes e que o ganho do desempenho paralelo é significativo à medida que mais processos MPI são utilizados.

O trabalho [Rosso and Schepke 2013] faz uma avaliação do desempenho do refinamento dinâmico de malhas no modelo OLAM e considera variações no refinamento dinâmico das malhas durante a simulação. Desse modo, faz testes de desempenho considerando um número distinto de processos MPI. [Osthoff et al. 2011] propôs a melhoria do desempenho do OLAM através de uma implementação híbrida com OpenMP e MPI. O trabalho observou que com o aumento do número de nós no cluster, a versão híbrida executa melhor que a versão em MPI. A principal razão é que a versão híbrida do OLAM diminui o número de arquivos de saída, resultando em um melhor desempenho para operações de E/S.

3. OLAM (*Ocean-Land-Atmosphere-Model*)

O modelo OLAM é um modelo de simulação numérica baseado no RAMS [Pielke 1992]. Ele estende as capacidades do RAMS para modelagem de um domínio global e tira proveito de várias de suas características como: parametrizações, métodos de inicialização, assimilação dos dados, estrutura lógica e, do código e formatos de entrada e saída [Walko and Avissar 2008]. O modelo pode representar fenômenos de escala global e regional simultaneamente, além de possuir processos dinâmicos de configuração da grade, estrutura de memória e técnicas de solução numérica [Silva et al. 2009]. Além disso, o OLAM utiliza técnicas de discretização de volumes finitos [Marshall et al. 1997] que resolvem as equações de Navier-Stokes.

A estrutura do modelo OLAM consiste em uma malha não estruturada de células triangulares. Estas células são dispostas horizontalmente sobre o globo terrestre e são constituídas de camadas verticais. A malha tem um formato inicial de icosaedro, formado de 20 triângulos, que engloba toda a esfera terrestre. Cada um desses triângulos pode ainda ser dividido, melhorando a resolução horizontal da malha. Como a malha não é estruturada, os triângulos não tem sua distribuição horizontal bem definida e suas coordenadas são armazenadas computacionalmente para serem posteriormente utilizadas na integração numérica [Silva et al. 2009].

A principal característica do modelo OLAM é a capacidade de simulação de fenômenos em escala global e local. Para tanto, o modelo utiliza um refinamento dinâmico

das malhas [Walko and Avissar 2008]. Esse refinamento baseia-se na subdivisão dos triângulos que compõe a região desejada em mais triângulos, aumentando a resolução dessa região. Quanto menor for a distância horizontal entre dois triângulos, maior é a resolução do modelo e mais dados são computados.

4. Avaliações e Testes Realizados

Os testes realizados buscaram evidenciar o custo da comunicação das rotinas MPI na execução do modelo. Para tanto, foram encontradas todas as rotinas de *send/receive* da implementação. Para mensurar o tempo de comunicação, foi aplicada a função *cpu time* do FORTRAN, em cada uma dessas rotinas. Foram realizadas 20 execuções com 1, 2, 4 e 8 processos, respectivamente. Na configuração do modelo, foi considerado um tempo de simulação equivalente a 1 hora e a resolução horizontal foi definida em 100 km.

A realização dos testes se deu em uma *workstation* Dell Precision T7600. A arquitetura possui 128 GB de memória e dois processadores Intel Xeon E5/Core i7, de 2.00 GHz. Cada processador possui 8 núcleos físicos e estendidos a 16 núcleos pelo uso da tecnologia *Hyper-Threading*. No total, são 32 núcleos de processamento, 16 físicos e 16 lógicos. A máquina roda Ubuntu/Linux 14.04 e contém instalações do modelo OLAM 3.0 e Open MPI 1.6.5.

5. Resultados Obtidos

No gráfico demonstrado na Figura 1, são comparados os tempos de comunicação dos processos MPI e computação do modelo OLAM. Há três barras distintas: a barra azul representa o processo de *rank* A, em que o tempo de comunicação é mínimo; a barra laranja que representa a média do tempo dos demais processos em execução, denominando-os de *rank* N e a barra amarela representa o tempo de computação.

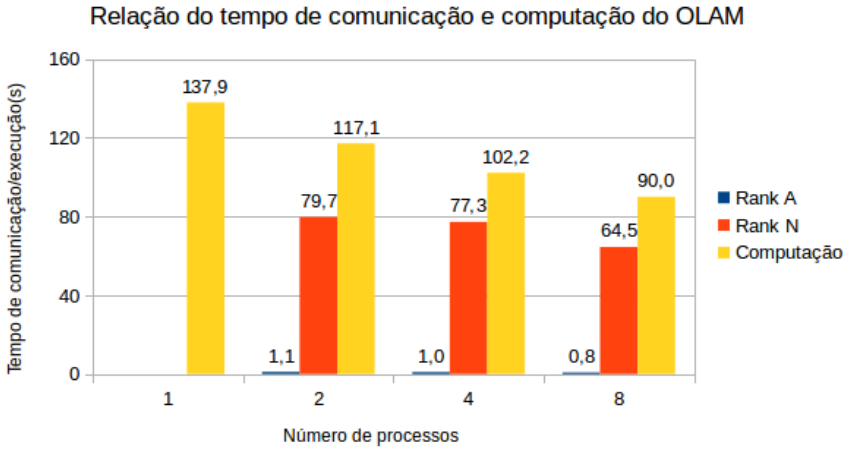

Figura 1. Relação entre as comunicações MPI e o tempo de execução do modelo OLAM

Os tempos de comunicação nos processos de *rank* A são ínfimos. Isso indica que esses processos não estão sendo sincronizados entre si e por consequência, apresentam tempos diferentes dos demais processos. O grupo de processos de *rank* N, tem o tempo

de comunicação mais elevado que os processos de *rank* A, apesar de serem mensurados de maneira idêntica. Esse fator supõe que ocorram chamadas de rotinas MPI assíncronas à execução. Além disso, é possível verificar que o aumento do número de processos traz benefícios tanto na computação quanto na comunicação do modelo. Entretanto, ampliar em excesso o número de processos pode fazer com que o tempo de computação seja prejudicado.

6. Conclusões e Trabalhos Futuros

Neste trabalho foi apresentada uma avaliação do tempo gasto para a comunicação de rotinas MPI implementadas no OLAM. Os resultados supõem que as chamadas MPI assíncronas mascaram o tempo de comunicação das rotinas de envio e recebimento. Adicionalmente, é possível verificar que os processos não são sincronizados entre si. Nesse caso, o tempo de comunicação do processo de *rank* A será diferente do tempo dos demais processos.

Em trabalhos futuros, pretende-se avaliar a comunicação do modelo em um sistema de memória distribuída e realizar um comparativo com as taxas de comunicação apuradas nesse trabalho. Além disso, será realizado um estudo mais minucioso com a finalidade de traçar um perfil da ocorrência das chamadas MPI no modelo OLAM.

Referências

Marshall, J., Adcroft, A., Hill, C., Perelman, L., and Heisey, C. (1997). A finite volume, incompressible navier stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research Oceans*, 102:5753–5766.

Osthoff, C., Grunmann, P., Boito, F., Kassick, R., Pilla, L., Navaux, P., Schepke, C., Panetta, J., Maillard, N., Silva Dias, P., and Walko, R. (2011). Improving Performance on Atmospheric Models through a Hybrid OpenMP/MPI Implementation. *Parallel and Distributed Processing with Applications (ISPA), 2011 IEEE 9th International Symposium on*, pages 69–74.

Pielke, R. (1992). A comprehensive meteorological model system - rams. *Meteorology and Atmospheric Physics*, 49:69–91.

Pydd, E., Karlinski, T., and Schepke, C. (2014). Avaliação de desempenho paralelo do modelo olam usando processos mpi em arquitetura de memória compartilhada. *Escola Regional de Alto Desempenho/RS*, pages 169–172.

Rosso, J.and Vargas, F. and Schepke, C. (2013). Avaliação de desempenho do refinamento dinâmico de malhas em modelo climatológico global. *XIV Simpósio em Sistemas Computacionais WSCAD-IC*, pages 254–256.

Schepke, C., Osthoff, C., Maillard, N., and Dias, P. (2010). Atmospheric simulation model performance evaluation on multicore environments. *Latin-American Conference on High Performance Computing*, pages 330–332.

Silva, R., Dias, P., Moreira, D., and Souza, E. (2009). Modelo OLAM (Ocean-Land-Atmosphere-Model): Descrição, Aplicações, e perspectivas. *Revista Brasileira de Meteorologia*, pages 144–157.

Walko, R. and Avissar, R. (2008). The ocean land atmosphere model (olam). part i: Shallow-water tests. *American Meteorological Society*, 136:4033–4044.