

Aplicação de Método Guloso para Balanceamento de Carga em uma Aplicação de Exploração Eletromagnética

Jessica Imlau Dagostini, Vinicius Garcia Pinto, Lucas Mello Schnorr

¹ Instituto de Informatica – Universidade Federal do Rio Grande do Sul (UFRGS)

{jidagostini, vgpinto, schnorr}@inf.ufrgs.br

Resumo. *O presente trabalho apresenta a aplicação de um método guloso para balanceamento de carga de uma aplicação de inversão de dados eletromagnéticos, coletados empregando o método mCSEM. Tal política trouxe um ganho de desempenho de 41% para a aplicação em relação à política original.*

1. Introdução

Técnicas de mapeamento eletromagnético como o *marine Controlled-Source Electromagnetic Method* (mCSEM) [Key 2009] são uma alternativa aos métodos sísmicos na busca por petróleo e gás no oceano. No mCSEM, um barco carrega um emissor eletromagnético que move-se sobre receptores fixados no fundo do mar. Quando estimulados pelo emissor, tais receptores capturam dados eletromagnéticos da área em questão. Estes dados permitem obter informações sobre o subsolo marinho através de inversão de dados. Neste trabalho, investigamos o comportamento de uma aplicação paralela baseada em MPI que realiza a inversão de dados coletados pelo mCSEM. Nesta análise, identificamos um desbalanceamento de carga e demonstramos uma solução gulosa para melhorar o desempenho da aplicação.

Para realizar a inversão dos dados coletados, os sinais são mapeados com a técnica de *Common Midpoints* (CMP). Nessa técnica, cada medida é mapeada para um ponto médio entre emissor e receptor. Diversos pares de Emissor-Receptor (E-R) são agregados para uma coleção de CMPs [Corrêa and Régis 2017], que são divididos entre os trabalhadores para serem processados pela aplicação. Neste processamento, há três regiões de execução: encaminhamento, correção e do cálculo de Cholesky.

Em sua versão original, a aplicação faz a divisão de carga distribuindo uma quantidade idêntica de CMPs para cada trabalhador. Todavia, cada CMP possui uma quantidade díspar de pares de E-R associados ao seu ponto. Em nossas análises, pudemos perceber que tal diferença entre os dados agregados a cada CMP é o responsável pelo desbalanceamento de carga encontrado na região de encaminhamento, uma vez que a carga é a informação associada a cada par E-R, que então é associado aos CMPs. Assim, identificamos que o desbalanceamento de carga encontrado na aplicação é diretamente relacionado a forma como ela realiza a divisão de trabalho entre os nós, pois ela não considera a quantidade de pares E-R agregados a cada CMP para tal repartição.

2. Solução, Resultados e Conclusão

Para mitigar o problema, nossa solução considerou a quantidade de pares E-R associados a cada CMP para realizar a distribuição entre os trabalhadores. Para isso, aplicamos uma estratégia gulosa, na qual se considera a carga já associada ao trabalhador para realizar novas atribuições. Apesar de ser mais custoso computacionalmente, tal estratégia

é uma solução ótima para problemas de particionamento numérico [Korf 2009]. Assim, alocamos um novo CMP a cada trabalhador respeitando a carga já atribuída a ele.

Em nossos resultados experimentais, a utilização do método guloso trouxe ganhos de desempenho comparado a estratégia original. Para este resultado, executamos a aplicação com 112 trabalhadores (número máximo de *cores* físicos disponíveis) em um cluster com 7 nós computacionais, cada um deles equipados com 2x Intel Xeon E5-2640v2 e 64GB de memória RAM. Nesta execução, forçamos a aplicação a realizar 100 iterações do método. Usando Score-P para realizar a instrumentação e coleta dos resultados, focamos nossa visualização para a região mais custosa da aplicação (encaminhamento). A Figura 1 apresenta uma visualização temporal da execução mostrando esta região. Cada trabalhador é identificado no eixo Y, enquanto o tempo de execução é representado no eixo X. O principal resultado é a redução de 41% no tempo de execução com a política gulosa ($\approx 4642s$) quando comparada à política original ($\approx 7832s$). Como a execução demonstrada têm repetições internas representativas (100 repetições das iterações), não houve necessidade de repetições do experimento.

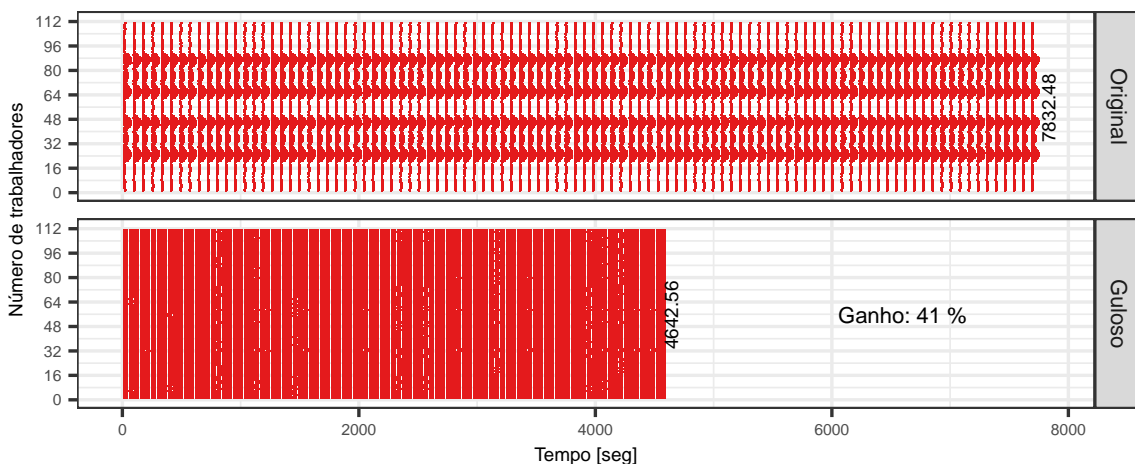


Figura 1. Comparação de desempenho entre as políticas Original e Gulosa.

Provamos assim que a abordagem gulosa trouxe ganhos de desempenho para a aplicação. Trabalhos futuros, agora com foco no uso de memória para que possamos escalar a execução da aplicação, já estão no mapa de melhorias.

Agradecimentos: trabalho parcialmente financiado por Petrobras (2018/00263-5). Experimentos executados no PCAD INF/UFRGS (<http://gppd-hpc.inf.ufrgs.br>).

Referências

- Corrêa, J. and Régis, C. T. (2017). 1d CMP inversion of MCSEM data to create a 3d geoelectrical model. In *SEG Technical Program Expanded Abstracts 2017*. Society of Exploration Geophysicists.
- Key, K. (2009). 1D inversion of multicomponent, multifrequency marine CSEM data: Methodology and synthetic studies for resolving thin resistive layers. *Geophysics*, 74(2):F9–F20.
- Korf, R. E. (2009). Multi-way number partitioning. In *Proc. 21st Int. Joint Conf. Artificial Intelligence*, page 538–543, San Francisco, USA. Morgan Kaufmann Publishers Inc.