

CLUSTER PARA REAMOSTRAGEM DE IMAGENS DO SATÉLITE GOES-12

Hugo J. Corrá¹, Fernando Ramos Martins¹, Marcio Magini²

¹INPE/DMA, Caixa Postal 515 - CEP 12245-970 – São José dos Campos, hugocorra@dge.inpe.br

INPE/DMA, Caixa Postal 515 - CEP 12245-970 – São José dos Campos, fernando@dge.inpe.br

²UNIVAP/FCC, Av. Shishima Hifumi, 2911 – CEP 12244-000 – S. José dos Campos, magini@univap.br

Resumo - O trabalho descreve a metodologia empregada para a reamostragem de imagens digitais do satélite meteorológico GOES. A metodologia emprega a técnica de processamento em cluster a fim de reduzir o elevado tempo de processamento para a reamostragem de imagens com 3241 colunas e 1820 linhas. Os testes finais estão sendo realizados e verificou-se que a técnica de cluster possibilitou a realização das reamostragens de forma a atender a demanda computacional. Testes finais estão sendo realizados a fim de determinar a redução do tempo de processamento com o emprego da técnica de cluster.)

Palavras-chave: cluster, sensoriamento remoto, processamento de imagens.

Área do Conhecimento: I – Ciências Exatas e da Terra

Introdução

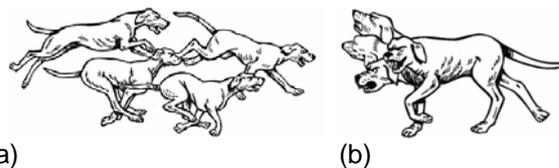
As novas teorias e descobertas científicas que podem ser compreendidas e assimiladas rapidamente são fonte de progresso tecnológico imediato. Porém, devido a grande quantidade de informações existentes nos dias atuais torna-se complicado conseguir compreendê-las e manipulá-las de forma rápida, conveniente e útil.

Modelos computacionais para modelagem, simulação e análise de dados são desenvolvidos e aprimorados diariamente, visando fornecer resultados passíveis de interpretação humana, ou visando gerar novos dados que sirvam de entrada para outros modelos. A medida em que a quantidade de informações cresce, cresce também a demanda por poder de processamento, o que gera a necessidade do desenvolvimento de soluções tecnológicas capazes de suprir esta demanda. Muitas vezes são criadas soluções muito poderosas e inovadoras, como é o caso dos super computadores, que fornecem um poder incrível de processamento, chegando a executar cerca de 70 trilhões de operações de ponto flutuante por segundo. Porém máquinas deste porte possuem um custo muito elevado, tornando esta tecnologia inviável para a resolução de muitos dos problemas existentes atualmente.

Cluster pode ser considerado como sendo uma coleção de sistemas computacionais conectados em rede, provendo uma série de recursos paralelos para aplicações e serviços de forma transparente a seus usuários, que terão a impressão que somente um único sistema computacional foi responsável pela execução do serviço.

As principais vantagens da utilização de clusters estão na construção simplificada com

hardware convencional; expansão e atualização facilitada em comparação com sistemas SMP (computador com dois ou mais processadores e memória compartilhada); custo mais baixo que grandes sistemas SMP com o mesmo número de CPUs; e permissão de uma escala flexível de performance. Podem ser consideradas como desvantagens; o aumento de complexidade, conhecimento especializado para construir e operar e a utilização de aplicações específicas para clusters. A Figura 1 procura ilustrar a diferença de um cluster para um sistema SMP.



(a) (b)
Figura 1 – Diferença entre um sistema de cluster e um sistema SMP; (a) Sistema de cluster, união de vários elementos para formar um sistema. (b) Sistema SMP, união de vários dos principais elementos para formar um sistema.

Materiais e Métodos

Imagens do satélite geoestacionário GOES-12

O cluster será desenvolvido para reamostrar imagens do satélite geoestacionário GOES-12, fornecidas pela Divisão de Satélites e Sistemas Ambientais – DSA, pertencente ao Centro de Previsão e Estudos Climáticos – CPTEC. As imagens são constituídas de 4 arquivos, 1 contendo os valores medidos pelo sensor, 2 arquivos de navegação contendo valores de latitude e longitude para cada ponto e 1 arquivo de

cabeçalho contendo informações a respeito da imagem.

A maior parte das imagens do satélite GOES-12 para a região delimitada pela América Latina (28,59° N / 110,99° W – 52,32° S / 5,69° W) possui cerca de 3421 pontos de largura por 1820 de altura, o que resulta em cerca de 6.226.220 (seis milhões, duzentos e vinte e seis mil, duzentos e vinte) elementos. Isso faz com que o processo de reamostragem tenha que executar milhões de verificações e interpolações para cada imagem a ser reamostrada.

Metodologia para reamostragem das imagens

Para cada pixel de uma imagem georeferenciada existe um valor de latitude e um de longitude. Nas imagens do satélite GOES-12, os valores de latitude e longitude variam para cada imagem, o que gera um problema quando se deseja comparar valores de pixel de uma mesma localidade em imagens diferentes. O pixel de uma imagem no horário de 12:15 pode possuir um valor de latitude e longitude diferente do mesmo pixel em uma imagem no horário de 12:45, isto é, uma localidade definida por suas coordenadas de latitude e longitude pode ser representada por pixels diferentes em imagens consecutivas. O processo de reamostragem visa padronizar valores de latitude e longitude para várias imagens, como forma de simplificar a interpretação e manipulação e comparação de imagens de satélite obtidas em horários e dias diferentes.

O processo de reamostragem é iniciado através da definição de uma navegação que será utilizada para todas as imagens reamostradas. Os valores da radiância medida pelo satélite para cada pixel da imagem reamostrada é determinado utilizando uma técnica de interpolação dos valores de radiância presentes na imagem original recebida do satélite. Para o processo de interpolação é necessário a definição de um raio, utilizado para especificar um intervalo de busca dos pixels localizados na proximidade das coordenadas (latitude e longitude) do pixel a ser preenchido na imagem reamostrada. Todos os pixels da imagem original do satélite que estão no intervalo entre o valor da latitude menos o raio até o valor da latitude mais o raio; e entre o valor da longitude menos o raio e o valor da longitude mais o raio, serão utilizados no processo matemático de interpolação para a obtenção do valor do pixel na imagem reamostrada. Para intervalos que não contenham nenhum pixel da imagem original do satélite, o valor no pixel reamostrado deve ser um código de erro (-32768).

A fim de reduzir o tempo de processamento necessário e atender a alta demanda computacional exigida no processo de

reamostragem de cada imagem do satélite GOES, optou-se pela aplicação da técnica de processamento em cluster. Conforme ilustrado na Figura 2, o princípio básico de funcionamento do cluster é o padrão de “divisão e conquista” [4], que propõe a divisão de um problema em subproblemas.

No caso do processo de reamostragem, a imagem que se deseja reamostrar deve ser dividida em blocos e distribuída entre os diversos nodos do cluster. Desta forma, o nodo principal (mestre) fica encarregado de dividir a imagem, enviar para os demais nodos (escravos) e esperar que os mesmos retornem os resultados do processamento. Por motivos de desempenho, o nodo principal é o computador onde as imagens estão armazenadas, pois desta forma o mesmo tem acesso direto às imagens, e pode carregá-las e salvá-las sem a necessidade de trafegar dados pela rede.

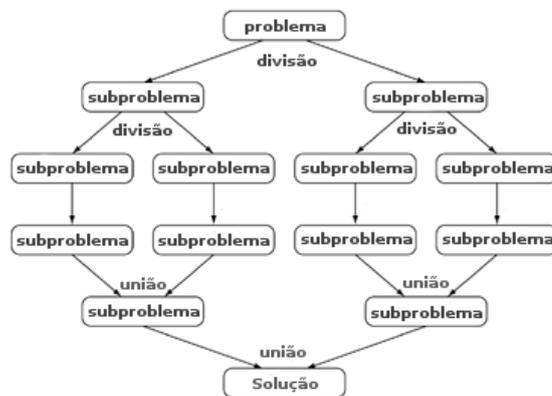


Figura 2 – Padrão de divisão e conquista propõe dividir um problema em problemas menores.

O cluster para reamostragem de imagens de satélite está sendo construído utilizando computadores com arquitetura x86, com 1 ou 2 processadores, sistema operacional GNU/Linux, C++ como linguagem de programação e MPI como mecanismo para troca de mensagens entre os nodos do cluster.

A maior dificuldade no desenvolvimento do cluster é fazer com que o software de reamostragem consiga enviar e receber de forma correta os dados e a otimização de mecanismos de busca por vizinho mais próximo. O balanceamento de carga também é muito importante, pois evita que um computador com menos recursos receba a mesma quantidade de trabalho que outros computadores com mais recursos. No momento de inicialização do cluster, o nodo mestre efetua uma verificação para identificar quais nodos estão respondendo mais rápido e quais possuem mais recursos disponíveis. Quando o software de reamostragem é iniciado o usuário especifica quantos processos o cluster

deve criar, e o cluster realiza a distribuição desses processos entre os nodos. Um nodo pode executar mais de um processo, dependendo de sua disponibilidade. Então, o software de reamostragem no nodo principal entra em um loop de envio e recebimento de dados, enquanto os demais nodos entram em loops, recebendo, processando e enviando de volta para o nodo principal os dados processados.

Resultados

O projeto ainda se encontra em fase de desenvolvimento, e não apresenta resultados conclusivos, porém é esperado que o sistema consiga realizar a reamostragem das imagens de forma a reduzir o elevado tempo de processamento para cada imagem de satélite.

Conclusão

Clusters foram inicialmente projetados para serem uma alternativa aos sistemas de super computação, que tem um custo muito elevado, não sendo acessíveis a grande parte das instituições e empresas. Através do desenvolvimento do sistema descrito no decorrer do trabalho, pretende-se obter poder de processamento suficiente para atender a alta demanda computacional requerida no processo de reamostragem das imagens do satélite GOES-12.

Referências

[1] MARTINS, F. R. Influência do processo de determinação da cobertura de nuvens e dos aerossóis de queimada no modelo físico de radiação BRASIL-SR. PhD. Thesis. INPE (Brazilian Institute of Space Research). São José dos Campos, Brasil, 2001.

[2] Martins, F. R., Souza-Echer, M. P., Pereira, E. B. Levantamento dos recursos de energia solar no Brasil com o emprego de satélite geoestacionário – o projeto swera. Revista Brasileira de Ensino de Física, v. 26, n2., p. 145-159, 2004.

[3] BAKER, M. R., BUYYA R. Cluster Computing: The Commodity Supercomputing - PRACTICE AND EXPERIENCE, VOL. 1(1), 1-4 (Janeiro 1998).

[4] Sanders, B., Massingil, B., Mattson, T. The Algorithm Structure Design Space in Parallel Programming. Disponível em: <http://www.informit.com/articles/article.asp?p=366887&rl=1>. Acesso dia 05/08/2005.