

Metodologia para determinação de climatologia de mesoescala do modelo BRAMS em Grade Computacional

*Eugenio Sper de Almeida
Centro de Previsão de Tempo e Estudos
Climáticos
Instituto Nacional de Pesquisas Espaciais
eugenio@cptec.inpe.br*

*Haroldo Fraga de Campos Velho
Laboratório de Computação Aplicada
Instituto Nacional de Pesquisas Espaciais
haroldo@lac.inpe.br*

*Airam Jonatas Preto
Laboratório de Computação Aplicada
Instituto Nacional de Pesquisas Espaciais
airam@lac.inpe.br*

*Stephan Stephany
Laboratório de Computação Aplicada
Instituto Nacional de Pesquisas Espaciais
stephan @lac.inpe.br*

Resumo

Este trabalho tem como objetivo apresentar uma metodologia para determinação de climatologia de mesoescala em grade computacional, utilizando o modelo meteorológico BRAMS - Brazilian Regional Atmospheric Modeling System. A geração da climatologia para um dado número de anos tem um alto custo computacional. Uma forma de acelerar o tempo de execução de uma climatologia é utilizar clusters de computadores que estão distribuídos em várias instituições de pesquisa. Para isto é necessário configurá-los para que constituam uma grade computacional, que é uma infra-estrutura de hardware e software que provê acesso confiável e consistente a recursos computacionais de alto desempenho que estão distribuídos fisicamente: neste caso, os clusters serão instalados no Departamento de Informática da Universidade Federal do Rio Grande do Sul (UFRGS), do Laboratório de Computação Aplicada (LAC) e Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), os dois últimos pertencentes ao Instituto Nacional de Pesquisas Espaciais (INPE) e localizados em São José dos Campos e Cachoeira Paulista respectivamente. Para validar esta metodologia, este trabalho apresenta uma simulação para a determinação desta a climatologia para os meses de janeiro, fevereiro e março de 2001 e 2002.

Palavras-chave: meteorologia, grade computacional, climatologia, modelos numéricos, BRAMS.

1. Introdução

Atualmente é possível simular as condições da atmosfera com até quinze (15) dias de antecedência e com acerto em até cinco (5) dias. Para isto são utilizados modelos numéricos de previsão de tempo (MNT). As condições da atmosfera podem ser determinadas em escala global ou regional. Para isto, são necessários computadores que possam executar esses modelos em um período curto de tempo, de modo que a informação seja disponibilizada antes que o evento meteorológico ocorra.

O Centro de Previsão de Tempo e Estudos Climáticos (CPTEC) do Instituto Nacional de Pesquisas Espaciais (INPE) utiliza supercomputação massivamente para a execução dos modelos globais, regionais, ensemble (a previsão de tempo por ensemble consiste em produzir condições atmosféricas iniciais perturbadas e rodar por diversas vezes o mesmo modelo a partir destas condições iniciais perturbadas - Mendonça e Bonatti, 2002), de onda e de clima, seja para pesquisa ou rotina operacional.

Dos modelos citados anteriormente, o modelo de clima é o que apresenta maior custo computacional, pois é um modelo que prevê

com antecedência o comportamento médio da atmosfera para um determinado período de meses. No entanto, para que os modelos de clima apresentem resultados confiáveis, é necessário que a climatologia gerada por esses modelos para o período considerado seja conhecida. Para isto é necessário que esse modelo execute a integração das informações referentes às condições da atmosfera por um longo período de anos, ou seja, o tempo de processamento de uma climatologia de 10 anos utiliza 520 mais espaço em disco do que uma previsão de tempo para sete (7) dias.

Para a determinação da climatologia será utilizada a tecnologia de computação em grade, que difere da computação distribuída convencional devido ao fato de ser voltada para compartilhamento de recursos em larga escala (como supercomputadores, “clusters”, sistema de armazenamento, dados, instrumentos e pessoas), aplicações inovadoras e em alguns casos, orientada ao alto desempenho (Foster et al., 2001).

Segundo Foster e Kesselman (1999), a palavra grade foi escolhida em analogia com as redes de energia elétrica (“electric power grid”). A idéia básica é disponibilizar recursos computacionais da mesma forma que hoje é disponibilizada a energia elétrica. Para viabilizar o uso de grades computacionais são necessárias redes de alta velocidade interconectando as instituições.

Uma descrição de modelos numéricos de previsão de tempo, com ênfase no modelo BRAMS, é apresentada na seção 2. A seção 3 descreve a tecnologia de grades computacionais, que é baseada em protocolos, e esta disponível no “Globus Toolkit”, que é disponibilizado dentro do escopo do “The Globus Project”. Nesta seção são apresentadas as ferramentas principais para a criação de grades computacionais.

Na seção 4 é descrita a grade computacional a ser utilizada, baseada em clusters de computadores com sistema operacional Linux e apresentada a metodologia para determinação da climatologia de mesoescala para o modelo BRAMS. Na seção 5 é

apresenta os resultados preliminares e na seção 6 as conclusões.

2. Modelos numéricos de previsão de clima

Nos dias atuais, previsão de tempo está fortemente relacionada com a análise dos resultados gerados por modelos numéricos de previsão de tempo, que são programas complexos que representam o movimento e os processos físicos da atmosfera através de equações matemáticas. Esses modelos recebem como parâmetros de entrada dados observacionais, dados derivados de imagens de satélite e dados gerados por modelos de dias anteriores. O processamento desses dados, utilizando-se as equações embutidas dentro dos modelos, gera arquivos de previsão numérica de tempo. Esses arquivos são representados na forma de matrizes, sendo cada matriz relacionada a uma determinada variável física, a um nível atmosférico e a um instante de tempo. Atualmente o CPTEC possui um modelo para previsão de tempo global - MCGA (Modelo de Circulação Geral Atmosférico) - e outro para previsão de tempo regional - Eta (η). O MCGA em uso no CPTEC é proveniente da versão 1.7 do COLA (“Center for Ocean, Land and Atmosphere Studies”) e é denominado versão CPTEC/COLA (Bonatti, 1996). O modelo Eta é um modelo de mesoescala de equações primitivas, que utiliza o método de ponto de grade (Chou, 1996). Estes modelos também são utilizados para previsão de clima, com as devidas modificações.

Outro modelo de área limitada que vem sendo utilizado no CPTEC/INPE é o BRAMS (“Brazilian Regional Atmospheric Modeling System”), projeto financiado pela FINEP (Financiadora de Estudos e Projetos) na área de computação de alto desempenho em Meteorologia. Este projeto visa melhorar o modelo RAMS (“Regional Atmospheric Modeling System”) nos quesitos documentação, programação e adaptação às condições brasileiras. Segundo Tremback e Walko (1997), o RAMS foi desenvolvido pelo

Departamento de Ciência Atmosférica da Universidade do Estado do Colorado a partir da junção dos modelos: CSU de nuvens/mesoescala (Tripoli e Cotton, 1982), versão hidrostática do modelo de nuvens (Tremback, 1990) e do modelo de brisa do mar (Mahrer e Pielke, 1977). Para Pielke et al. (1992), dentre todas as melhorias introduzidas no RAMS com a junção desses modelos, a capacidade de aninhamento de grades é a mais importante, pois permite a criação de uma grade detalhada (maior resolução) a partir de uma grade de menor resolução (é possível aninhar mais de uma grade dentro de uma grade de menor resolução). O RAMS possui uma série de características e opções (discutidas detalhadamente por Walko e Tremback, 1991) que podem ser configuradas em tempo de execução.

Para simulações de previsão de tempo, o RAMS necessita de dados de prognósticos para as fronteiras laterais (condição de contorno) provenientes de um modelo global e de dados de análise para as condições iniciais. Esta fronteira altera seus valores a cada passo de tempo através de uma média ponderada entre os valores iniciais e finais das saídas do modelo global para uma dada previsão do modelo RAMS. A cada passo de tempo o RAMS calcula a previsão de tempo forçando os dados de sua fronteira com os dados da média ponderada do modelo global, através de uma técnica de assimilação de dados (“nudging”).

O resultado de uma previsão, gerada por um modelo meteorológico para um dado instante de tempo, é armazenado em um conjunto de matrizes tridimensionais (cada uma representando uma variável). Esses campos também podem ser vistos como colunas verticais sobre pontos de grade.

O código do BRAMS permite execução paralela, para isto o conjunto de pontos de grade que compõem a análise é transformado em subconjuntos retangulares de pontos de grade (subdomínios). Nesta estrutura de subdomínios, segundo Tremback e Walko (1997), emprega-se o modelo mestre-escravo.

Neste modelo, um processo mestre controla a execução dos demais processos (escravos), que são efetivamente os responsáveis pela execução das tarefas. No BRAMS, o processo mestre manipula a inicialização e a saída e processos escravos executam a computação do modelo.

Cada nó escravo recebe um subdomínio, com a respectiva região de fronteira, e executa a computação. A cada passo de tempo (“timestep”), os nós trocam as informações sobre as fronteiras dos subdomínios, utilizando a biblioteca de passagem de mensagens MPI (Gropp et al., 1999) e o processamento continua até a geração de todos os campos para os horários previstos.

3. Computação em grade

Segundo Foster et al. (2001), a Computação em Grade difere da computação distribuída convencional devido ao fato de possibilitar compartilhamento coordenado de recursos em larga escala (como supercomputadores, “clusters”, sistema de armazenamento, dados, instrumentos e pessoas) e permitir a resolução de problemas computacionais em organizações virtuais multi-institucionais, em alguns casos orientados ao alto desempenho. O compartilhamento deve necessariamente ser altamente controlado, com os fornecedores de recursos e consumidores definindo claramente e cuidadosamente apenas o que é compartilhado, quem é permitido compartilhar e as condições nas qual o compartilhamento ocorre. Para atingir este objetivo a palavra chave é interoperabilidade, que em ambiente de rede significa protocolos comuns. Desta forma, uma arquitetura de grade nada mais é que uma arquitetura de protocolos.

Atualmente não existe um padrão internacional para tecnologia de Computação em Grade, mas o “Globus Toolkit” tem se tornado um padrão “de facto”. Ele foi desenvolvido dentro do escopo do “The Globus Project” que visa desenvolver

tecnologia para a criação de Grades Computacionais, que são ambientes que permitem a integração de softwares com instrumentos, displays, recursos computacionais e de informação que são gerenciados por diversas organizações espalhadas em várias localizações.

Os principais componentes do “Globus Toolkit” incluem “Grid Security Infrastructure” (GSI), que fornece serviços de autenticação baseada em chave pública e autorização local; serviços de gerenciamento de recursos, que fornece uma linguagem para especificação de requisitos de aplicação, mecanismos para reserva imediata e posterior dos recursos da grade e gerenciamento remoto de tarefas; e serviços de informação que recupera e distribui informações sobre os recursos da grade. Os serviços de grade de dados complementam e ampliam esses componentes: serviço de transferência GridFTP e serviço de gerenciamento de réplica.

4. Metodologia

A climatologia de mesoescala será determinada a partir de execuções consecutivas do modelo BRAMS em clusters instalados no Departamento de Informática da Universidade Federal do Rio Grande do Sul (UFRGS), do Laboratório de Computação Aplicada (LAC) e Centro de Previsão de Tempo e Estudos Climáticos (CPTEC), os dois últimos pertencentes ao Instituto Nacional de Pesquisas Espaciais (INPE) e localizados em São José dos Campos e Cachoeira Paulista respectivamente.

A metodologia a ser adotada prevê que a climatologia de mesoescala do modelo BRAMS de um determinado período seja executado em um dado cluster. Para isto, o modelo BRAMS será integrado para o dado período e ao fim da execução do mesmo é calculada a climatologia intermediária daquele período (Figura 4.1).

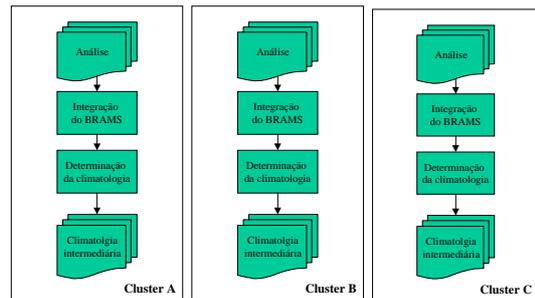


Figura 4.1 – Determinação da climatologia intermediária.

Após a execução de todos os períodos, os dados de climatologia intermediária resultantes serão transferidos para o cluster principal, que determinará a climatologia final dos períodos considerados (Figura 4.2).

A área escolhida para testar esta metodologia esta delimitada pelas coordenadas geográficas das longitudes 54° O e 45° O e das latitudes 19° S e 27° S, utilizando uma resolução de 111x111 km.

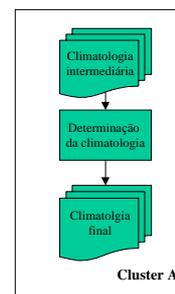


Figura 4.2 – Determinação da climatologia final.

Será utilizado o “Globus Toolkit” em todos os clusters. Esta ferramenta permite a execução remota de programas em computadores localizados em diferentes localizações geográficas.

Como o objetivo inicial será executar o modelo BRAMS na grade (apenas S.O. Linux), muitos dos serviços para construção de grades (fornecidos pelo “Globus Toolkit”) não serão implementados e utilizados, como suporte a diretórios agregados e gerenciamento e escalonamento de processos. De acordo com a necessidade, os serviços para construção de grades serão implementados.

5. Resultados preliminares

A metodologia apresentada foi validada em um cluster composto de 10 processadores Intel/Xeon de 3.0 GHz para os meses de janeiro, fevereiro e março de 2001 e 2003.

Os programas para determinação da climatologia intermediária foram executados, para cada ano considerado, no mesmo cluster em instantes de tempo diferentes. Ao final da execução dos mesmos, simulou-se a transferência dos dados gerados para o cluster principal que executou o programa que gera a climatologia final. Os dados foram validados de forma qualitativa com dados de climatologia do CPTEC/INPE referente a esses meses. A Figura 5.1 apresenta a climatologia gerada para esses meses.

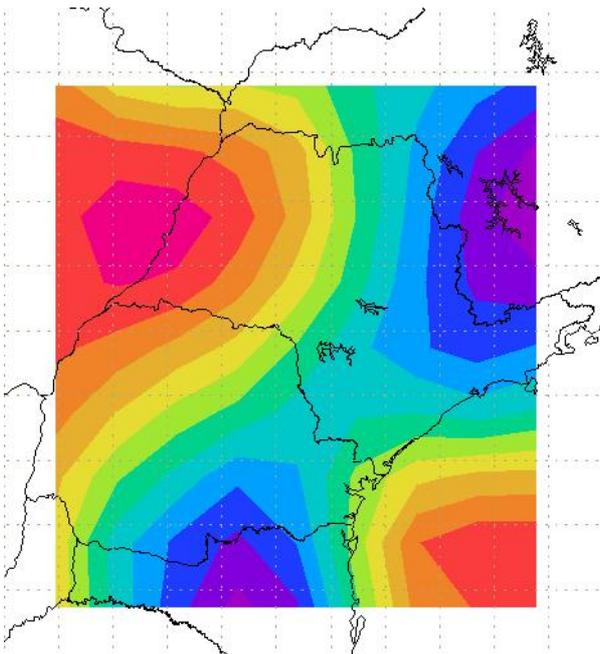


Figura 4.2 – Climatologia do BRAMS para o mês de janeiro.

6. Conclusões

Este trabalho apresentou a metodologia para determinação da climatologia de mesoescala do modelo BRAMS para execução em grade computacional, composta de clusters de computadores. Através dos resultados preliminares foi possível validar esta metodologia. O próximo passo será

montar a grade computacional nas três entidades envolvidas (LAC-INPE, CPTEC/INPE e UFRGS) para a validação final da metodologia.

Computadores com poder computacional para resolver problemas na área de meteorologia não são uma realidade em instituições brasileiras. Agregar computadores com relativo poder computacional de algumas instituições, como no caso de clusters de computadores, pode ser uma alternativa para resolver este problema. Desta forma, espera-se que trabalho possa contribuir para que num futuro pesquisadores possam executar aplicações meteorológicas em uma grade computacional nacional de Meteorologia. Dentro deste enfoque, o modelo meteorológico BRAMS é bastante adequado devido ao fato que ele pode ser executado em grades computacionais, composta “clusters” de máquinas escalares.

6. Referências

- Bonatti, J.P. **Modelo de circulação geral atmosférico do CPTEC.** CLIMANÁLISE: Boletim de monitoramento e análise climática, edição especial comemorativa de 10 anos. <http://www.cptec.inpe.br/products/climanalise/cliesp10a/bonatti.html>, 1996.
- Chou, S.C. **Modelo Regional ETA.** CLIMANÁLISE: Boletim de monitoramento e análise climática, edição especial comemorativa de 10 anos. <http://www.cptec.inpe.br/products/climanalise/cliesp10a/27.html>, 1996.
- Foster, I.; Kesselman, C.; Tuecke, S. **The Anatomy of the Grid: Enabling Scalable Virtual Organizations.** International J. Supercomputer Applications, 15(3), 2001.
- Foster, I.; Kesselman, C. **The Grid: Blueprint for a New Computing Infrastructure.** Morgan Kaufman Publishers, Inc., 1999.
- Gropp, W.; Lusk, E.; Skjellum, A.; **Using MPI: Portable Parallel Programming with the Message-Passing Interface.** 2nd edition, MIT Press, 1999.

- Mahrer, Y., Pielke, R.A. **A numerical study of the airflow over irregular terrain.** Beitr. Phys. Atmos., 50, 98-113, 1977.
- Mendonça, A.M.; Bonatti, J.P. **O sistema de previsão de tempo por ensemble do CPTEC.** Anais do XII Congresso Brasileiro de Meteorologia, Foz do Iguaçu, PR, agosto de 2002.
- Pielke, R.A.; Cotton, W.R.; Walko, R.L.; Tremback, C.J.; Lyons, W.A.; Grasso, L.D. Nicholls, M.E.; Moran, M.D.; Wesley, D.A.; Lee, T.J.; Copeland J.H. **A comprehensive meteorological modeling system - RAMS.** Meteor. Atmos. Phys., 49, 69-91, 1992.
- Tripoli, G.J.; Cotton, W.R. **The Colorado State University three-dimensional cloud mesoscale model, 1982: Part I: General theoretical framework and sensitivity experiments.** J. de Rech. Atmos., 16, 185-220, 1982.
- Tremback, C.J.; Walko, R.L. **The Regional Atmospheric Modeling System (RAMS): Development of parallel processing computer architectures.** 3rd RAMS Users Workshop, Echuca, Victoria, Australia, July 1997
- Tremback, C.J. **Numerical simulation of a mesoscale convective complex model development and numerical results.** Ph.D. dissertation, Atmos. Sci. Paper No. 465, Department of Atmospheric Science, Colorado State University, Fort Collins, CO 80523, 247pp., 1990.
- Walko, R.L.; Tremback, C.J. **RAMS - The Regional Atmospheric Modeling System Version 2C: User's guide.** Published by ASTeR, Inc., P.O. Box 466, Fort Collins, Colorado. 86pp, 1991.