

# Análise de desempenho de um sistema de modelagem atmosférica em nuvens computacionais

Mateus S. de Melo<sup>1</sup>, Lúcia M. A. Drummond<sup>1</sup>, Roberto P. Souto<sup>2</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal Fluminense (UFF)  
Niterói, RJ

<sup>2</sup>Laboratório Nacional de Computação Científica  
Petrópolis, RJ

msmelo@id.uff.br, lucia@ic.uff.br, rpsouto@lncc.br

**Abstract.** *The use of cloud computing in high performance computing has raised questions about the advantage of this type of system over on-premises systems. However, comparing these two systems is not trivial, due to several factors such as the number of solutions that cloud computing providers offer and also the behavior of the application. This study aimed to analyze the numerical weather prediction model BRAMS in both systems. For a small case study, the application performs similarly across systems. The cost of running the application in different markets offered by AWS was also analyzed, which for the problem used, it is advisable to use the spot market.*

**Resumo.** *A utilização de nuvens computacionais em computação de alto desempenho tem levantado questões sobre a vantagem desse tipo de sistema em relação a sistemas tradicionais (on premise). Porém, comparar esses dois sistemas não é trivial, devido a diversos fatores como a quantidade de soluções que os provedores de nuvens oferecem e também do comportamento da aplicação. Este estudo visou analisar o modelo numérico de previsão de tempo e clima BRAMS nos dois sistemas. Observou-se que para um estudo de caso pequeno, a aplicação possui desempenho semelhantemente nos sistemas. Também foi analisado o custo da execução da aplicação em diferentes mercados ofertados pela AWS, que para o problema utilizado, é aconselhado o uso do mercado spot.*

## 1. Introdução

Os supercomputadores são um tipo de sistema computacional comumente utilizado em computação de alto desempenho. Sendo muito utilizados na execução de aplicações científicas e simulações, os supercomputadores possuem um alto poder de processamento e memória. Contudo, a aquisição e manutenção desse tipo de sistema é economicamente custosa, dificultando a sua ampla posse por instituições de pesquisa [Montes et al. 2020]. Além disso, o ambiente é dividido entre vários usuários, resultando, frequentemente, em longos tempos de espera na fila de execução.

A nuvem computacional pública é um modelo onde os recursos de infraestrutura e serviços de computação são oferecidos por provedores, podendo ser acessados remotamente e contratados sob demanda. Ela tem sido muito utilizada em aplicações na indústria [Stieninger and Nedbal 2014] e vem surgindo nos últimos anos pesquisas sobre

como utilizá-las em computação de alto desempenho. Entretanto, conseguir o melhor custo-benefício nesse tipo de sistema comparado aos supercomputadores tradicionais não é uma tarefa trivial. Há vários tipos de ambientes oferecidos pelos provedores que influenciam no preço final e no desempenho da aplicação [Netto et al. 2018].

Pode-se encontrar no campo da meteorologia modelos numéricos de previsão de tempo e clima, nos quais se beneficiam com o uso da computação de alto desempenho [Michalakes 2020], pois há uma grande quantidade de dados a serem processados, em equações que vão sendo resolvidas a fim de atualizar o avanço do estado da atmosfera. Entretanto, de nosso conhecimento, ainda são poucos os trabalhos que executam essas aplicações em nuvens computacionais [Montes et al. 2020] [Powers et al. 2021].

O objetivo principal desta pesquisa é analisar e comparar diversos ambientes oferecidos pela nuvem computacional *Amazon Web Services* (AWS) e comparar com supercomputadores tradicionais (*on-premise*) no que diz respeito ao custo e desempenho. A aplicação que será utilizada será o BRAMS (*Brazilian developments on the Regional Atmospheric Modeling System*)<sup>1</sup>, um modelo numérico de previsão do tempo e clima de escala regional mantido pelo INPE/CPTEC (Instituto Nacional de Pesquisas Espaciais/Centro de Previsão de Tempo e Estudos Climáticos), baseado no modelo RAMS (*Regional Atmospheric Modeling System*) [Pielke et al. 1992].

## 2. Modelo regional BRAMS

O modelo regional *Brazilian developments on the Regional Atmospheric Modeling System* (BRAMS) [Freitas et al. 2017] foi desenvolvido principalmente pelo INPE/CPTEC, baseando-se no modelo *Regional Atmospheric Modeling System* (RAMS) [Pielke et al. 1992], lançado na década de 1980 pela *Colorado State University* nos Estados Unidos.

O BRAMS oferece um modelo especializado na previsão regional de tempo e de clima no Brasil [Freitas et al. 2016], diferenciando-se do RAMS sobretudo ao introduzir novas funcionalidades que melhor representam os fenômenos meteorológicos tropicais [Freitas et al. 2009]. Do ponto de vista computacional, foram implementadas alterações que visaram o paralelismo massivo do modelo. Como estratégia de paralelização, foi empregada a decomposição do domínio da previsão, onde a atualização da atmosfera é realizada em paralelo nas diferentes sub-regiões do domínio. A implementação desta estratégia foi realizada por meio da biblioteca de passagem de mensagens MPI [Gropp et al. 1996, Gropp et al. 1999], permitindo escalabilidade de 9600 núcleos computacionais [Fazenda et al. 2012] de acordo com os critérios operacionais do INPE/CPTEC, sendo possível rodar uma previsão de 5Km em 20 minutos de tempo de execução.

## 3. Estratégia

Com objetivo de comparar o desempenho do modelo BRAMS em diferentes sistemas computacionais, foi escolhido o *cluster* disponibilizado pela Intel via *Intel Developer Cloud*<sup>2</sup> como tipo de sistema *on premise* a ser avaliado e a nuvem computacional pública *Amazon Web Services* (AWS).

---

<sup>1</sup>Código-fonte disponível em: <https://github.com/luflarois/brams>

<sup>2</sup><https://www.intel.com/content/www/us/en/developer/tools/devcloud/overview.html>

O serviço *Intel Developer Cloud*, ou Intel DevCloud, oferece um *cluster* com diversos recursos de hardware e software da Intel para o aprendizado de desenvolvimento de aplicações usando as soluções da Intel e também para avaliação do comportamento de aplicações nesse sistema. Para ter como referência o comportamento da aplicação em ambiente *on premise*, foi escolhida uma fila focada em ambiente *multicore* com processadores mais potentes entre as disponíveis, Intel Xeon Gold 6128. Esses processadores são da geração *Skylake* e a fila possui 192 GB de memória principal.

Já a *Amazon Web Service (AWS)* é uma plataforma de nuvem computacional pública amplamente adotada em todo mundo, fornecendo diversos tipos de serviços computacionais, entre eles, a oferta de máquinas virtuais escaláveis e com diferentes configurações. Neste trabalho foi escolhido executar o modelo na instância *z1d.6xlarge*. Embora esse tipo de instância conte com processadores Intel Xeon Platinum da série 8000, enquanto o nó executado na Intel DevCloud conta com processadores Intel Gold da série 6000, essa instância foi a que se aproximou melhor das configurações do nó ofertado pela Intel, levando em consideração a capacidade da memória principal e a geração de processadores *Skylake*. Além disso, os testes na nuvem foram separados em dois grupos distintos, sendo um grupo utilizando instâncias *on demand*, onde a máquina virtual é reservada especialmente para o usuário, e outro grupo usando instâncias *spot*, que são máquinas virtuais que usam a capacidade sobressalente disponível por um menor preço, mas que podem ser interrompidas a qualquer momento.

#### 4. Resultados

Os testes foram realizados com a previsão configurada com a resolução espacial de 20Km x 20Km e com a grade de 60 x 60 pontos, com 45 níveis verticais. O tempo de previsão foi de 24 horas com *timestep* de 120 segundos, totalizando 720 *timesteps*, e com o cálculo do esquema da radiação na atmosfera sendo feito a cada 1.800 segundos, totalizando 48 *timesteps*, 6,67% do total de *timesteps*. O modelo foi executado nas máquinas utilizando 1, 2, 4, 8, 12 processos MPI com uma execução cada e comparado seus resultados analisando o tempo total de execução, tempo total dos *timesteps* sem o cálculo da radiação e tempo total de *timesteps* com a radiação ativada. O nó utilizado na plataforma Intel DevCloud contavam com 2 processadores *Skylake Intel Xeon Gold 6128* com 6 núcleos cada, totalizando 12 núcleos no nó, memória principal de 192 GB e foi utilizado compiladores da GNU versão 11.

A tabela 1 apresenta os tempos de execução, tempos do total de *timesteps* sem o cálculo da radiação e os tempos do total de *timesteps* com o cálculo da radiação das execuções feitas na Intel DevCloud. Percebe-se que os *timesteps* que realizam o cálculo da radiação levaram quase o mesmo tempo daqueles que não realizam o cálculo, mesmo que em menor quantidade. Também é possível perceber que houve uma redução de tempo significativa quando foi adicionado mais processos MPI, sendo mostrado também na tabela 1 que o ganho de desempenho da aplicação foi sempre próximo do ideal, com os *timesteps* com a radiação ativada tendo um ganho de desempenho ligeiramente maior do que os que a radiação estava desativada.

Já na plataforma AWS, foi escolhida a instância *z1d.6xlarge* que possui processadores *Skylake Intel Xeon Platinum 8151* com 24 núcleos por instância, 192GB de memória principal e também foi utilizado compiladores da GNU versão 11. Além disso, foram

**Tabela 1. Tempos de execução e *Speedup* do modelo na Intel DevCloud**

Processos	Tempo	Tempo	Tempo	Média	Média	<i>Speedup</i>	<i>Speedup</i>	<i>Speedup</i>
	Total (s)	<i>Timestep</i> (s)	Radiação (s)	<i>Timestep</i> (s)	Radiação (s)	Total	<i>Timestep</i>	Radiação
1	711,8	366,625	343,250	0,500	7,000	1,00	1,00	1,00
2	373,9	203,250	169,125	0,250	3,500	1,90	1,80	2,03
4	200,2	110,375	88,750	0,125	1,750	3,56	3,32	3,87
8	111,8	65,467	45,500	0,125	0,875	6,37	5,60	7,54
12	78,8	44,334	33,132	0,062	0,688	9,03	8,27	10,36

escolhidos os mercados *on demand* e *spot* para fazer os testes. Como pode ser visto na tabela 2, os tempos obtidos na AWS escolhendo o mercado *on demand* foram um pouco menores que os tempos na Intel DevCloud, mas ainda próximos. Tendo o tempo de execução do modelo executado com 12 processos alcançado 73,3 segundos na AWS *on demand* e 78,8 segundos na Intel DevCloud. Consequentemente, a aplicação continuou atingindo um bom ganho de desempenho na AWS, como pode ser visto na tabela 2. Os tempos alcançados no mercado *spot* da AWS também foram muito semelhantes ao mercado *on demand*. Foi cobrado no mercado *on demand* \$2,23 por hora de uso da máquina. Já no mercado *spot*, foi cobrado \$1,00 por hora de uso. Através desses valores, é possível calcular o custo de cada da execução, como é visto na tabela 2.

**Tabela 2. Tempos de execução, *Speedup* e Custo do BRAMS na AWS**

Processos	Tempo	Tempo	<i>Timestep</i>	Média	Média	<i>Speedup</i>	<i>Speedup</i>	<i>Speedup</i>	Custo	Custo
	Total (s)	<i>Timestep</i> (s)	Radiação (s)	<i>Timestep</i> (s)	Radiação (s)	Total	<i>Timestep</i>	Radiação	<i>On Demand</i>	<i>Spot</i>
1	694,2	354,025	339,724	0,504	7,029	1,00	1,00	1,00	\$0,43	\$0,20
2	355,6	186,009	169,183	0,264	3,493	1,95	1,90	2,01	\$0,22	\$0,10
4	186,4	100,953	85,058	0,141	1,760	3,72	3,51	3,99	\$0,12	\$0,05
8	103,0	57,246	45,388	0,080	0,935	6,74	6,18	7,48	\$0,06	\$0,03
12	73,3	41,350	31,582	0,057	0,650	9,47	8,56	10,76	\$0,05	\$0,02

## 5. Conclusão

Ao estudar nuvens computacionais voltadas para computação de alto desempenho, foi visto que a execução de modelos numéricos de previsão de tempo e clima nesses sistemas tem ganhado seu espaço, mas ainda sendo pouco explorado, principalmente a respeito do modelo BRAMS. O presente trabalho tratou-se de um estudo introdutório, consistindo em utilizar dados menores que uma previsão tradicional para analisar o comportamento do modelo BRAMS dentre diferentes sistemas, *on premise* e em nuvens computacionais, sendo executados apenas em um nó. O desempenho do modelo para esse estudo de caso foi quase semelhante nos dois sistemas, sendo as execuções na AWS não sendo afetadas pela virtualização da máquina. Em todos os testes realizados nos dois sistemas, o modelo conseguiu alcançar um ganho de desempenho sempre próximo do ideal. O custo das execuções do modelo foram baixas, especialmente no mercado *spot*, que conseguiu ser a metade do valor do mercado *on demand*. O baixo custo aliado ao fato do problema ser pequeno ao ponto da execução não ser interrompida, faz com que o *spot* seja priorizado.

Como trabalhos futuros é pretendido executar o modelo nos dois sistemas com dados do tamanho de uma previsão tradicional. Além disso, é planejado incluir também testes de execuções em múltiplos nós computacionais.

## Agradecimentos

Este trabalho foi desenvolvido com recursos disponibilizados através do programa Intel Developer Cloud e também pelo Projeto Universal/CNPq número 404087/2021-3.

## Referências

- Fazenda, A. L., Rodrigues, E. R., Tomita, S. S., Panetta, J., and Mendes, C. L. (2012). Improving the scalability of an operational scientific application in a large multi-core cluster. In *Computer Systems (WSCAD-SSC), 2012 13th Symposium on*, pages 126–132. IEEE.
- Freitas, S., Longo, K., Silva Dias, M., Chatfield, R., Silva Dias, P., Artaxo, P., Andreae, M., Grell, G., Rodrigues, L., Fazenda, A., et al. (2009). The coupled aerosol and tracer transport model to the brazilian developments on the regional atmospheric modeling system (catt-brams)—part 1: Model description and evaluation. *Atmospheric Chemistry and Physics*, 9(8):2843–2861.
- Freitas, S. R., Panetta, J., Longo, K. M., Rodrigues, L. F., Moreira, D. S., Rosario, N. E., Silva Dias, P. L., Silva Dias, M. A., Souza, E. P., Freitas, E. D., et al. (2017). The brazilian developments on the regional atmospheric modeling system (brams 5.2): an integrated environmental model tuned for tropical areas. *Geoscientific Model Development*, 10(1):189–222.
- Freitas, S. R., Rodrigues, L. F., Panetta, J., Longo, K., Moreira, D., Freitas, E., Longo, M., Fazenda, A., Fonseca, R., Stockler, R., and Camponogara, G. (2016). *Description of the model input namelist parameters*. CPTEC/INPE, São Paulo, Brasil.
- Gropp, W., Lusk, E., Doss, N., and Skjellum, A. (1996). A high-performance, portable implementation of the mpi message passing interface standard. *Parallel computing*, 22(6):789–828.
- Gropp, W., Lusk, E., and Skjellum, A. (1999). *Using MPI: portable parallel programming with the message-passing interface*, volume 1. MIT press.
- Michalakes, J. (2020). Hpc for weather forecasting. *Parallel Algorithms in Computational Science and Engineering*, pages 297–323.
- Montes, D., Añel, J. A., Wallom, D. C., Uhe, P., Caderno, P. V., and Pena, T. F. (2020). Cloud computing for climate modelling: Evaluation, challenges and benefits. *Computers*, 9(2):52.
- Netto, M. A., Calheiros, R. N., Rodrigues, E. R., Cunha, R. L., and Buyya, R. (2018). Hpc cloud for scientific and business applications: taxonomy, vision, and research challenges. *ACM Computing Surveys (CSUR)*, 51(1):1–29.
- Pielke, R. A., Cotton, W., Walko, R. e. a., Tremback, C. J., Lyons, W. A., Grasso, L., Nicholls, M., Moran, M., Wesley, D., Lee, T., et al. (1992). A comprehensive meteorological modeling system—rams. *Meteorology and Atmospheric Physics*, 49(1-4):69–91.
- Powers, J. G., Werner, K. K., Gill, D. O., Lin, Y.-L., and Schumacher, R. S. (2021). Cloud computing efforts for the weather research and forecasting model. *Bulletin of the American Meteorological Society*, 102(6):E1261–E1274.
- Stieninger, M. and Nedbal, D. (2014). Diffusion and acceptance of cloud computing in smes: towards a valence model of relevant factors. In *2014 47th Hawaii International Conference on System Sciences*, pages 3307–3316. IEEE.