

Avaliação reprodutível de políticas de alocação e migração de VMs em *data centers* geodistribuídos sustentáveis

Guilherme Fernandes Moraes da Silva, Daniel Cordeiro

¹Escola de Artes, Ciências e Humanidades – Universidade de São Paulo (USP)
São Paulo, SP – Brasil

{guilherme.fernandes01,daniel.cordeiro}@usp.br

Abstract. Comparing virtual machine (VM) management policies is not trivial because each study adopts a different simulator, workloads, and metrics, compromising comparability across studies. This work presents a standardized, reproducible evaluation of 10 VM allocation and migration policies (four of them carbon-aware) in geo-distributed data centers, using SimGrid: under a single environment (9 data centers, Google 2011 and Azure 2020 workloads, power model, carbon series, and metrics), 200 simulations were run (10 policies $\times$ 2 workloads $\times$ 10 seeds). The same carbon-aware policy yielded the lowest emissions under both workloads, the ranking of the remaining policies changed with the workload, and lower energy consumption did not always imply lower emissions: the form of evaluation directly influences the conclusions.

Resumo. Comparar políticas de gerenciamento de máquinas virtuais (VMs) não é trivial, pois cada estudo adota um simulador, cargas de trabalho e métricas distintos, comprometendo a comparabilidade entre estudos. Este trabalho apresenta uma avaliação padronizada e reprodutível de 10 políticas de alocação e migração de VMs (quatro delas cientes de carbono) em *data centers* geodistribuídos, usando o SimGrid: no mesmo ambiente (9 *data centers*, cargas Google 2011 e Azure 2020, modelo de potência, séries de carbono e métricas), foram realizadas 200 execuções (10 políticas $\times$ 2 cargas $\times$ 10 sementes). Nas duas cargas, a mesma política ciente de carbono obteve as menores emissões, mas o ranking das demais variou com a carga e menor consumo de energia nem sempre implicou menores emissões: a forma de avaliação influencia as conclusões.

1. Introdução

Data centers sustentam serviços em nuvem, aplicações distribuídas e sistemas de inteligência artificial. Estima-se que já consumam cerca de 2% da eletricidade global e respondam, junto às redes de transmissão de dados, por aproximadamente 1% das emissões de gases de efeito estufa associadas à energia [Masanet et al. 2020]; a *International Energy Agency* (IEA) projeta que a expansão dos *data centers* responderá por cerca de metade do crescimento de mais de 420 TWh na demanda elétrica dos Estados Unidos até 2030; na China, a demanda do subsetor de serviços digitais, que inclui *data centers* e redes, cresceu 17% em 2025.¹ De forma complementar, a *U.S. Energy Information Administration* (EIA) projeta que o consumo elétrico apenas dos servidores de *data centers* nos Estados Unidos pode atingir entre 446 e 818 bilhões de kWh até 2050.²

¹<https://www.iea.org/reports/electricity-2026>

²<https://www.eia.gov/outlooks/aeo/narrative/>

Como *data centers* geodistribuídos têm acesso a diferentes matrizes energéticas, algoritmos de escalonamento cientes de carbono (*carbon-aware*) podem ser utilizados para mitigar essas emissões [Vasconcelos et al. 2023]. No entanto, comparar esses algoritmos de forma justa exige cautela metodológica, pois o resultado depende do ambiente experimental adotado.

Diante desse contexto, este artigo propõe uma avaliação padronizada e reproduzível no SimGrid [Casanova et al. 2025], na qual 10 políticas de gerenciamento de VMs, oito de alocação inicial e duas que também realizam migrações, são submetidas ao mesmo ambiente, às mesmas cargas, séries regionais de carbono e métricas, para responder se o desempenho relativo das políticas muda com a carga, se minimizar energia minimiza emissões e qual é o efeito das migrações. O trabalho estende o estudo da ERAD-SP 2025 [Silva and Cordeiro 2025] para 9 *data centers* geodistribuídos.

2. Conceitos fundamentais

Vasconcelos et al. [Vasconcelos et al. 2022] investigaram estratégias cientes de carbono, como o impacto indireto da rede em abordagens *follow-the-renewables*, que motivam parte das políticas e do ambiente deste artigo. Como avaliar em ambientes reais é oneroso, simuladores são a plataforma usual para essa comparação. Contudo, a literatura costuma variar em simulador, topologia, carga, modelo de potência e séries de carbono, o que dificulta determinar se as diferenças decorrem dos algoritmos ou do ambiente experimental. Este trabalho preenche essa lacuna ao padronizar o ambiente de avaliação, em vez de propor mais um algoritmo.

3. Metodologia

Os experimentos usam o SimGrid sobre 9 *data centers* geodistribuídos (450 *hosts*, 50 por *data center* em três grupos heterogêneos), baseados em [Vasconcelos et al. 2022], com *Power Usage Effectiveness* (PUE) por região. Cada região (Canberra, Seul, Paris, Virgínia, Dubai, Singapura, Pune, Joanesburgo e São Paulo) tem uma série horária de intensidade de carbono ($\text{g CO}_2/\text{kWh}$) de 2020. Foram empregados dois traços reais, Google 2011 e Azure 2020, idênticos para todas as políticas: chegadas e demandas vêm dos traços, a semente sorteia apenas os tempos de vida das VMs e os demais parâmetros são fixos. As emissões integram, a cada minuto, a energia de cada *data center* (ponderada pelo PUE) vezes a intensidade de carbono horária local.

Quatro das 10 políticas avaliadas são cientes de carbono: *wsnb* [Xu and Buyya 2020], *followme_s* (FollowMe@S) [Ali et al. 2021], *cnemesis* (*c-nemesis*) [Vasconcelos et al. 2022] e *lowest_carbon_dc*, gulosa ingênua que aloca sempre no *data center* de menor intensidade de carbono. As outras seis são agnósticas a carbono: as heurísticas clássicas *best_fit*, *bfd* (*best-fit decreasing*), *ffd* (*first-fit decreasing*), *first_fit* e *worst_fit*, além do *round_robin*. Apenas *cnemesis* e *followme_s* também realizam migrações periódicas (a cada 10 min, até 20 e 50 por decisão), intra e entre *data centers*; as demais atuam só na alocação inicial. As migrações usam a *live migration* do SimGrid (pré-cópia com reenvio de páginas modificadas): a transferência ocupa a rede e mantém origem e destino ligados, custo capturado no consumo; a energia de rede e a degradação das aplicações, porém, não são modeladas. O experimento cruza 10 algoritmos $\times$ 2 cargas $\times$ 10 sementes

(200 execuções). As métricas são consumo de energia, emissões, número de migrações e volume transferido. Código, dados de entrada e instruções de reprodução estão publicamente disponíveis.³

4. Resultados e discussão

As políticas cientes de carbono reduzem as emissões (Tabela 1) e, nas duas cargas avaliadas, a menor emissão foi da `lowest_carbon_dc`: a redução frente à pior política (`round_robin`) chega a cerca de 96% no Google 2011 e 91% no Azure 2020 e, frente à `best_fit`, fica em torno de 81% e 75%. Duas cargas, porém, não bastam para concluir superioridade geral; a carga altera a magnitude das diferenças e a posição relativa das demais políticas: o `followme_s` tem a 4ª menor emissão no Google 2011, mas cai para a 6ª no Azure 2020, superado por `ffd` e `first_fit`. Comparar algoritmos com uma única carga, ou cargas distintas entre estudos, pode levar a conclusões equivocadas.

Tabela 1. Média (desvio-padrão) sobre 10 sementes, ordenada por emissões no Google 2011.

Algoritmo	Google 2011		Azure 2020	
	kWh	gCO ₂	kWh	gCO ₂
<i>Cientes de carbono</i>				
<code>lowest_carbon_dc</code>	27,2 (0,5)	2.175 (81)	41,0 (0,6)	5.326 (164)
<code>cnemesis</code>	29,1 (0,5)	2.324 (86)	42,6 (0,4)	5.500 (124)
<code>wsnb</code>	30,4 (0,6)	5.046 (172)	41,0 (0,3)	5.981 (93)
<code>followme_s</code>	112,4 (11,2)	9.199 (1.011)	113,5 (11,9)	14.786 (1.553)
<i>Agnósticas a carbono (geo-blind)</i>				
<code>ffd</code>	28,0 (0,3)	10.137 (72)	45,5 (0,2)	13.199 (58)
<code>first_fit</code>	31,5 (0,8)	11.024 (213)	45,2 (0,5)	13.161 (64)
<code>bfd</code>	31,2 (0,3)	11.434 (125)	47,1 (0,6)	21.382 (220)
<code>best_fit</code>	31,1 (0,6)	11.452 (205)	45,1 (0,5)	21.366 (219)
<code>worst_fit</code>	68,4 (0,0)	31.359 (22)	65,3 (1,0)	29.986 (512)
<code>round_robin</code>	139,3 (1,6)	63.208 (988)	131,7 (1,4)	60.659 (1.216)

Menor consumo de energia não significa menores emissões: a `ffd` está entre as de menor consumo no Google 2011 (28,0 kWh), mas tem apenas a 5ª menor emissão (10.137 gCO₂), muito acima da `cnemesis` (2.324 gCO₂), que consome mais (29,1 kWh). Como as emissões são o produto da energia de cada *data center* pela intensidade de carbono local na hora do consumo, uma política pode gastar mais energia e emitir menos, se concentrar a carga em regiões e horários de eletricidade mais limpa. As heurísticas de consolidação ignoram essa dimensão; avaliar só o consumo subestima o ganho ambiental das políticas cientes de carbono.

Nas migrações, `cnemesis` e `followme_s` esgotam seus orçamentos em todas as decisões (460 e 1150 por execução; cerca de 4 e 10 TB transferidos). A `followme_s` consome de 2,7 a 4× mais energia que as demais cientes de carbono, coerente com sua movimentação contínua, que mantém mais *hosts* ligados e retransmite memória; é também a mais sensível à semente ($\pm 11,9$ kWh e ± 1.553 gCO₂, ante variação abaixo de

³<https://github.com/GuilhermeFernandes01/erad-sp-2026/tree/erad-sp-2026>

4% nas demais). Já a `cnemesis` obteve emissões próximas às da `lowest_carbon_dc`, que não migra: aqui, migrar não superou alocar bem desde o início.

5. Considerações finais

Foi apresentada uma avaliação padronizada e reproduzível de 10 políticas de alocação e migração de VMs (200 execuções no SimGrid sob o mesmo ambiente), cuja contribuição é metodológica: um protocolo de comparação justa que mostra como a forma de avaliação influencia os resultados. Nas duas cargas avaliadas, a `lowest_carbon_dc` obteve as menores emissões, o ordenamento das demais políticas variou com a carga, minimizar energia não implicou minimizar emissões e migrar não superou a melhor alocação inicial. As conclusões se restringem ao cenário avaliado: duas cargas e uma topologia fixa não permitem generalizar o vencedor, e o custo das migrações é modelado só parcialmente (sem rede nem impacto nas aplicações). Trabalhos futuros incluem ampliar cargas, topologias e análises estatísticas, e incorporar custos, SLA e latência.

Este trabalho foi realizado com apoio da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), procs. #2019/26702-8, #2021/06867-2 e #2023/00811-0, do CCD Cidades Carbono Neutro (FAPESP #2024/01115-0), do CNPq (proc. #444766/2024-3) e do grupo de pesquisa THUS (Techno-Human Systems of the Future) do Centro Internacional de Pesquisa (IRC) ‘Transitions’ (CNRS, USP e FAPESP CIP #2025/01171-0).

Referências

- Ali, H., Zakarya, M., Rahman, I. U., Khan, A. A., and Buyya, R. (2021). Followme@ls: Electricity price and source aware resource management in geographically distributed heterogeneous datacenters. *Journal of Systems and Software*, 175:110907.
- Casanova, H., Giersch, A., Legrand, A., Quinson, M., and Suter, F. (2025). Lowering entry barriers to developing custom simulators of distributed applications and platforms with SimGrid. *Parallel Computing*, 123:103–125.
- Masanet, E., Shehabi, A., Lei, N., Smith, S., and Koomey, J. (2020). Recalibrating global data center energy-use estimates. *Science*, 367(6481):984–986.
- Silva, G. and Cordeiro, D. (2025). Migração de máquinas virtuais como estratégia para data centers sustentáveis. In *Anais da XVI Escola Regional de Alto Desempenho de São Paulo*, pages 70–73, Porto Alegre, RS, Brasil. SBC.
- Vasconcelos, M., Cordeiro, D., Da Costa, G., Dufossé, F., Nicod, J.-M., and Rehn-Sonigo, V. (2023). Optimal sizing of a globally distributed low carbon cloud federation. In *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, pages 203–215. IEEE.
- Vasconcelos, M. F. S., Cordeiro, D., and Dufossé, F. (2022). Indirect network impact on the energy consumption in multi-clouds for follow-the-renewables approaches. In *Proceedings of the 11th International Conference on Smart Cities and Green ICT Systems (SMARTGREENS)*, pages 44–55.
- Xu, M. and Buyya, R. (2020). Managing renewable energy and carbon footprint in multi-cloud computing environments. *Journal of Parallel and Distributed Computing*, 135:191–202.