

Escalonamento com Consciência Energética para Fluxos de Trabalho Científicos sem Servidor: Uma Abordagem de Aprendizado de Máquina

Lucas Rosa¹, Alfredo Goldman¹

¹Instituto de Matemática e Estatística (IME), Departamento de Ciência da Computação
Universidade de São Paulo (USP) — São Paulo, SP — Brasil

{roses.lucas, gold}@ime.usp.br

Abstract. *This article proposes to address the challenges of energy efficiency and workflow scheduling in serverless computing environments. By integrating machine learning techniques and simulation, the research aims to bridge gaps between energy efficiency and serverless scheduling. The methodology involves historical data collection, energy consumption prediction through machine learning, and the development of scheduling policies with deep neural networks. The project also includes adaptation of workflow management systems and validation in real-world environments, aiming to provide viable solutions to current challenges in HPC.*

Resumo. *Este artigo propõe abordar os desafios de eficiência energética e escalonamento de fluxos de trabalho científicos em ambientes de computação sem servidor. Integrando técnicas de aprendizado de máquina e simulação, a pesquisa visa preencher lacunas entre eficiência energética e escalonamento sem servidor. A metodologia inclui coleta de dados históricos, previsão de consumo energético por meio de aprendizado de máquina e desenvolvimento de políticas de escalonamento com redes neurais profundas. O projeto também envolve adaptação de sistemas de gerenciamento de fluxo de trabalho e validação em ambientes reais, visando oferecer soluções viáveis para os desafios atuais em HPC.*

1. Motivação

Um dos grandes desafios contemporâneos na área da Computação de Alto Desempenho (HPC) é a busca pela eficiência energética. Os supercomputadores têm desempenhado um papel importante em uma ampla gama de áreas, incluindo pesquisa sobre mudanças climáticas, mitigação de desastres naturais, modelagem molecular, inteligência artificial, entre outras. Estima-se que alcançaremos a computação Zettascale – o próximo grande marco na área de HPC – na próxima década [Liao et al. 2018]. No entanto, a energia necessária para operar nessa capacidade será de meio Gigawatt [Su and Naffziger 2023], o que equivale a metade da capacidade elétrica de uma planta nuclear.

No âmbito do software, a eficiência energética pode ser alcançada através do aprimoramento da gestão de recursos e escalonamento de tarefas. No entanto, é um desafio [Majid and Marin 2023] gerenciar simultaneamente vários objetivos e equilibrar as compensações entre custo, tempo de execução, consumo de energia e utilização de recursos.

Por outro lado, as aplicações científicas, geralmente modeladas como monólitos [Roy et al. 2022b], podem ser representadas por grafos direcionados acíclicos (DAGs). Sistemas de gerenciamento de fluxos de trabalho (WMSs), como Pegasus, Airflow e Argo, simplificam sua implementação ao converterem o DAG abstrato em listas de tarefas priorizadas, posteriormente escalonadas por gerenciadores de recursos como Kubernetes, HTCondor ou Slurm.

As aplicações científicas modernas apresentam uma complexidade inerente devido à interdependência dos dados e à execução de múltiplas tarefas em paralelo. Elas demandam uma crescente variedade de recursos computacionais, que incluem não apenas CPUs, mas também GPUs e FPGAs. No entanto, parte dessa complexidade não é comunicada do WMS para o gerenciador de recursos, levando a decisões de escalonamento não-ótimas [Lehmann et al. 2023]. Além disso, ao escalonar tais aplicações em clusters de HPC com alocação estática de recursos, é comum observar tanto subutilização quanto superutilização desses recursos [Roy et al. 2022b, Roy et al. 2022a].

Diante deste cenário, a computação sem servidor, uma tendência na computação em nuvem, oferece uma oportunidade promissora devido à sua capacidade de dimensionamento automático de recursos conforme a demanda, aliada a um modelo de cobrança flexível baseado no uso efetivo dos recursos. Apesar das vantagens, o escalonamento de fluxos de trabalho científicos sem servidor ainda apresenta desafios [Elshamy et al. 2023], como a execução sem estado, o tempo limite de execução e a sobrecarga de escalabilidade.

2. Proposta

A pesquisa neste artigo visa melhorar o escalonamento de fluxos de trabalho científicos sem servidor, integrando fatores de sustentabilidade como o consumo de energia em um processo de tomada de decisão multiobjetivo. Nosso objetivo é superar abordagens tradicionais, como heurísticas e meta-heurísticas, por meio do uso de técnicas de aprendizado de máquina para desenvolver agentes inteligentes responsáveis pela tomada de decisão. Dessa forma, nosso foco se encontra nos seguintes desafios de pesquisa.

DP1: *Integração do escalonamento de fluxos de trabalho científicos com ambientes de computação sem servidor, otimização multiobjetivo, recursos heterogêneos e algoritmos de aprendizado de máquina.*

Um dos objetivos desta pesquisa é a integração de vários aspectos no escalonamento de fluxo de trabalho científico em ambientes de computação sem servidor. Esses ambientes são notavelmente caracterizados por suas demandas em constante mudança e pela variedade de capacidades e recursos de processamento disponíveis.

Além disso, nos últimos anos, tem-se observado um aumento significativo na adoção de técnicas de aprendizado de máquina para otimizar tanto o gerenciamento de recursos quanto o escalonamento de tarefas [Zhang et al. 2020, Antici et al. 2023, Koslovski et al. 2024]. Os modelos de aprendizagem por reforço profundo (DRL) aprendem a tomar decisões interagindo com ambientes e ajustando estratégias para maximizar recompensas acumuladas. Esta habilidade de aprender e adaptar-se torna o DRL uma solução eficaz para o ambiente dinâmico e complexo mencionado.

DP2: *Falta de dados de consumo de energia e informações heterogêneas sobre recursos nos registros de fluxos de trabalho científicos.*

Trabalhar com aprendizado de máquina requer dados de treinamento confiáveis, frequentemente derivados de registros do mundo real ou de modelos realistas. Esses registros são vitais para pesquisas em escalonamento de tarefas.

Repositórios abertos, incluindo o Workflow Trace Archive¹ (WTA) e WfInstances², fornecem dados abrangentes de registros de execuções de fluxo de trabalho científico. No entanto, a ausência de dados de consumo de energia das tarefas em mais de 250 registros nos últimos quatro anos representa uma limitação significativa.

Além disso, os conjuntos de dados de ambos os repositórios não fornecem suporte claro para recursos heterogêneos e, como resultado, também não incluem dados relacionados à energia específicos para esses recursos. Portanto, adquirir registros ricos em informações é essencial para obter dados mais representativos.

DP3: *Criação de formas mais robustas para validação de estratégias de escalonamento de fluxos de trabalho científicos.*

Na literatura, incluindo os estudos referenciados [Zhang et al. 2020, Antici et al. 2023, Koslovski et al. 2024], é evidente a falta de experimentos validados em ambientes reais e a ausência de fluxos de trabalho que façam uso de recursos heterogêneos. No entanto, a realização de experimentos em ambientes reais, com cargas de trabalho realistas, é necessária para uma comparação precisa entre diversos sistemas e algoritmos.

3. Metodologia

Inicialmente, propomos resolver o desafio **DP2** obtendo dados detalhados sobre consumo de energia e características de recursos heterogêneos. Planejamos usar o Argo Workflows³ como *engine* de execução e estabelecer um *framework* para coleta de dados. Também pretendemos enriquecer os repositórios Workflow Trace Archive e WfInstances com registros contendo métricas de consumo de energia em nível de tarefa, podendo definir um formato expandido, se necessário.

Visando sanar o desafio **DP1** e com os dados coletados, empregaremos técnicas de aprendizado de máquina supervisionado, especialmente modelos de regressão, para modelar o consumo de energia das tarefas de fluxos de trabalho. Concomitantemente, pretendemos desenvolver políticas de escalonamento de fluxos de trabalho científicos por meio de um agente baseado em DRL e, se possível, avaliar a eficácia das técnicas de aprendizagem online e contínua na adaptação às mudanças ambientais em tempo real.

Por fim, para abordar o desafio **DP3**, planejamos adaptar o WMS Argo para dar suporte à nossa política de escalonamento, a fim de validar sua eficácia em um ambiente real. Isso incluirá a criação de um canal de comunicação entre o Argo e o Kubernetes, além de ajustes no escalonador do Kubernetes para integrar esses dados durante os processos de escalonamento. Essa configuração não apenas validará nossa política de escalonamento, mas também proporcionará a validação de outras estratégias desenvolvidas na literatura, as quais anteriormente eram testadas em ambientes pouco realistas.

¹<https://wta.atlarge-research.com/>

²<https://wfcommons.org/instances>

³<https://argoproj.github.io/workflows/>

Referências

- Antici, F., Yamamoto, K., Domke, J., and Kiziltan, Z. (2023). Augmenting ML-based Predictive Modelling with NLP to Forecast a Job's Power Consumption. In *Proceedings of the SC '23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis*, pages 1820–1830, Denver CO USA. ACM.
- Elshamy, A., Alquraan, A., and Al-Kiswany, S. (2023). A Study of Orchestration Approaches for Scientific Workflows in Serverless Computing. In *Proceedings of the 1st Workshop on SErverless Systems, Applications and MEdethodologies*, pages 34–40, Rome Italy. ACM.
- Koslovski, G. P., Pereira, K., and Albuquerque, P. R. (2024). DAG-based workflows scheduling using Actor–Critic Deep Reinforcement Learning. *Future Generation Computer Systems*, 150:354–363.
- Lehmann, F., Bader, J., Tschirpke, F., Thamsen, L., and Leser, U. (2023). How Workflow Engines Should Talk to Resource Managers: A Proposal for a Common Workflow Scheduling Interface. In *2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing (CCGrid)*, pages 166–179. arXiv:2302.07652 [cs].
- Liao, X.-k., Lu, K., Yang, C.-q., Li, J.-w., Yuan, Y., Lai, M.-c., Huang, L.-b., Lu, P.-j., Fang, J.-b., Ren, J., and Shen, J. (2018). Moving from exascale to zettascale computing: challenges and techniques. *Frontiers of Information Technology & Electronic Engineering*, 19(10):1236–1244.
- Majid, A. Y. and Marin, E. (2023). A Review of Deep Reinforcement Learning in Serverless Computing: Function Scheduling and Resource Auto-Scaling.
- Roy, R. B., Patel, T., Gadepally, V., and Tiwari, D. (2022a). Mashup: making serverless computing useful for HPC workflows via hybrid execution. In *Proceedings of the 27th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, pages 46–60, Seoul Republic of Korea. ACM.
- Roy, R. B., Patel, T., and Tiwari, D. (2022b). DayDream: Executing Dynamic Scientific Workflows on Serverless Platforms with Hot Starts. In *SC22: International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 1–18, Dallas, TX, USA. IEEE.
- Su, L. and Naffziger, S. (2023). 1.1 Innovation For the Next Decade of Compute Efficiency. In *2023 IEEE International Solid-State Circuits Conference (ISSCC)*, pages 8–12, San Francisco, CA, USA. IEEE.
- Zhang, D., Dai, D., He, Y., Bao, F. S., and Xie, B. (2020). RLScheduler: an automated HPC batch job scheduler using reinforcement learning. In *Proceedings of the international conference for high performance computing, networking, storage and analysis, Sc '20*. IEEE Press.