

Hierarchical Multi-Agent Framework for Carbon-Efficient Liquid-Cooled Data Center Clusters

Soumyendu Sarkar^{1*†}, Avisek Naug^{1*}, Antonio Guillen^{1*}, Vineet Gundecha^{1*},
Ricardo Luna Gutiérrez¹, Sahand Ghorbanpour¹, Sajad Mousavi¹, Ashwin Ramesh Babu¹,
Desik Rengarajan^{1,2}, Cullen Bash¹

¹Hewlett Packard Enterprise

²Amazon

{soumyendu.sarkar, avisek.naug, antonio.guillen, vineet.gundecha, rluna, sahand.ghorbanpour, sajad.mousavi, ashwin.ramesh-babu, desik.rengarajan, cullen.bash}@hpe.com

Abstract

Reducing the environmental impact of cloud computing requires efficient workload distribution across geographically dispersed Data Center Clusters (DCCs) and simultaneously optimizing liquid and air (HVAC) cooling with time shift of workloads within individual data centers (DC). This paper introduces Green-DCC, which proposes a Reinforcement Learning (RL) based hierarchical controller to optimize both workload and liquid cooling dynamically in a DCC. By incorporating factors such as weather, carbon intensity, and resource availability, Green-DCC addresses realistic constraints and interdependencies. We demonstrate how the system optimizes multiple data centers synchronously, enabling the scope of digital twins, and compare the performance of various RL approaches based on carbon emissions and sustainability metrics while also offering a framework and benchmark simulation for broader ML research in sustainability.

workload distribution and energy efficiency (Ahmad, Khalil, and Shah 2020; Guo et al. 2013; Chang, Roohi, and Gao 2019; Peng et al. 2017; Khalil, Ahmad, and Almazroi 2019). The challenges are at two levels: the individual data center level (low level) and the data center cluster level (top level). As discussed in our previous work on holistic real-time data center control (Sarkar et al. 2024b,a, 2023; Naug et al. 2024), there are complex internal dependencies between multiple controls, such as time shift of workload and cooling. In addition, dynamic external dependencies, such as workload, weather patterns, and carbon intensity, require an adaptive approach like multiagent reinforcement learning (MRL) that static solutions cannot provide. The top-level control for the data center cluster is a more difficult challenge. This is because while the top-level control attempts to optimally distribute the load between data centers, the low-level controller at the data center level time shifts the workload, which causes a difficult synchronization challenge.

Introduction

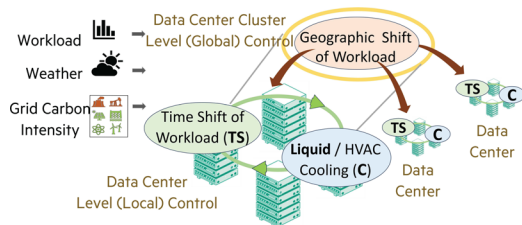


Figure 1: Spatio temporal hierarchical workload distribution.

To address these sustainability challenges with the exponential growth of AI and digital services, major cloud service providers with geographically distributed data centers are exploring innovative approaches to optimize

*These authors contributed equally.

†Corresponding Author

Proposed Method

The architecture employs a hierarchical RL framework to optimize workload distribution across geographically dispersed data centers by a global agent, followed by temporal rescheduling of the loads at each data center by a dedicated agent. Tiered decisions at the geographic and local data center levels are based on carbon intensity information and the workload assigned to each data center. Figure 1 il-

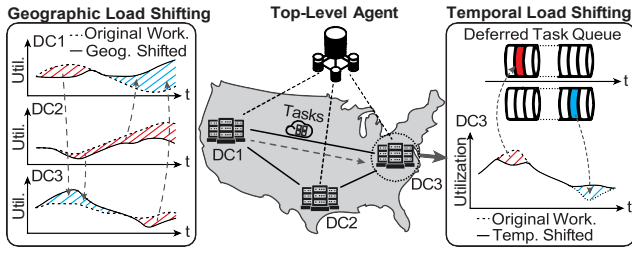


Figure 2: Left: Geographic load shifting between data centers (DC1, DC2, DC3). Right: Temporal load shifting within a data center.

illustrates the geographical load shifting strategy, where tasks are dynamically moved between data centers to optimize resource utilization and reduce environmental impact. Lower-level agents handle localized operations within each data center, focusing on temporal workload scheduling. Noncritical tasks are deferred using a Deferred Task Queue (DTQ) for processing during low grid carbon intensity time periods. This combined approach of geographic and temporal load shifting enables our approach to optimize data center operations’ sustainability and efficiency at both global and local levels. Figure 2 shows this overall spatiotemporal load shifting strategy. We also introduce a dynamic cost model for workload transfers between data centers that account for geographical distance, transfer capacity, and resource constraints, balancing carbon efficiency with effective resource utilization.

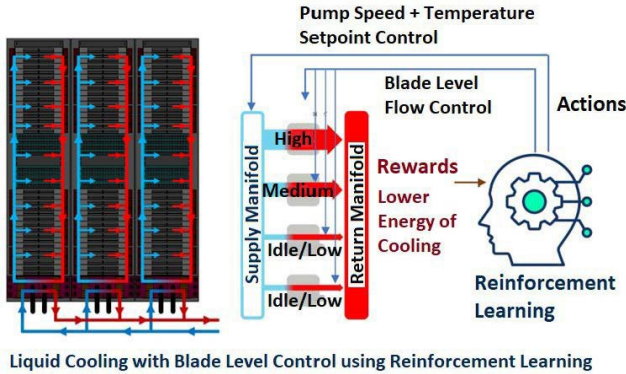


Figure 3: Blade Level RL control for Liquid Cooling

AI workload with GPUs, high performance computing, and high density servers, along with the sustainability mandate, have triggered growing adoption of high efficiency liquid cooled (LC) data centers. However, while data centers are making hardware improvements, effective control of liquid cooling is lacking, with most industry implementations using open-loop control. Our work proposes RL agents to dynamically control liquid flow and temperature setpoints. Additionally, this RL agent controls the liquid flow all the way to the blade level for the emerging LC data centers, as shown in Figure 3.

Overall, the combination of a top-level geographical load

shifting agent followed by a temporal load shifting and cooling optimizer agent at each data center in a hierarchical framework allows us to perform holistic optimization for data center clusters leveraging geographic differences in external parameters like weather and power grid carbon intensity. We provide a simulation framework to enable other ML researchers. The code, licenses, and instructions for using Green-DCC can be found on GitHub¹ along with the documentation².

Config	Agent			
	Baseline	Top RL	Top+Low RL	HRL
CO ₂	3845 ± 3	3515 ± 3	3512 ± 3	3435 ± 5

Table 1: Evaluation with load shift only with Hierarchical RL (HRL) showing best results.

Config.	Hierarchical RL Algorithm		
	PPO	A2C	APPO
CO ₂	3435 ± 5	3601 ± 3	3608 ± 3

Table 2: Tones of CO₂ produced by different RL algorithms. ± shows the standard deviation over the 10 seeds.

Results

We evaluated a data center cluster with locations in New York, California, and Georgia, each with unique grid carbon intensity, weather, and workload utilization patterns. As shown in Table 1, three variations of the RL algorithm were tested: (1) Top-Level RL, which trains a top-level agent for geographic workload distribution without lower-level temporal load shifting; (2) Top + Low-Level RL, combining geographic and temporal distribution with a pre-trained low-level agent; and (3) Hierarchical RL (HRL), where all agents are trained simultaneously.

PPO (Schulman et al. 2017) yielded the best results in all cases. Only load shifts were used for comparison. Although hierarchical RL is the most complex to train, it showed the best performance with PPO, minimizing the most CO₂ emissions. The results are presented in Tables 1 and 2.

Conclusion

Green-DCC represents a significant step forward in comprehensive spatio-temporal workload optimization in data center clusters using hierarchical reinforcement learning. By providing a realistic, comprehensive benchmark environment that captures the complexities of geographical and temporal workload shifting, we enable researchers and practitioners to develop innovative solutions that can dramatically reduce the carbon footprint of data center clusters.

¹GitHub repository: github.com/HewlettPackard/green-dcc.

²Documentation: hewlettpackard.github.io/green-dcc.

Acknowledgments

We would like to thank Paolo Faraboschi for sharing his expertise in machine learning and practical implementation approaches and Torsten Wilde for his feedback on energy optimization and sustainability.

Furthermore, we extend our gratitude to Wes Brewer, Feiyi Wang, Vineet Kumar, Scott Greenwood, Matthias Maiterth, and Terry Jones of Oak Ridge National Laboratory for their feedback and leadership within the ExaDigiT consortium, which helped refine our solution.

References

- Ahmad, I.; Khalil, M. I. K.; and Shah, S. A. A. 2020. Optimization-based workload distribution in geographically distributed data centers: A survey. *Int. J. Commun. Syst.*, 33(12): e4453.
- Chang, Y.-C.; Roohi, N.; and Gao, S. 2019. Neural lyapunov control. *Advances in neural information processing systems*, 32.
- Guo, Y.; Gong, Y.; Fang, Y.; Khargonekar, P. P.; and Geng, X. 2013. Energy and network aware workload management for sustainable data centers with thermal storage. *IEEE Transactions on Parallel and Distributed Systems*, 25(8): 2030–2042.
- Khalil, M. I. K.; Ahmad, I.; and Almazroi, A. A. 2019. Energy efficient indivisible workload distribution in geographically distributed data centers. *IEEE Access*, 7: 82672–82680.
- Naug, A.; Guillen, A.; Gutierrez, R. L.; Gundecha, V.; Bash, C.; Ghorbanpour, S.; Mousavi, S.; Babu, A. R.; Markovikj, D.; Kashyap, L. D.; Rengarajan, D.; and Sarkar, S. 2024. SustainDC: Benchmarking for Sustainable Data Center Control. In *The Thirty-eight Conference on Neural Information Processing Systems Datasets and Benchmarks Track*.
- Peng, Y.; Kang, D.-K.; Al-Hazemi, F.; and Youn, C.-H. 2017. Energy and QoS aware resource allocation for heterogeneous sustainable cloud datacenters. *Optical Switching and Networking*, 23: 225–240.
- Sarkar, S.; Naug, A.; Guillen, A.; Luna, R.; Gundecha, V.; Ramesh Babu, A.; and Mousavi, S. 2024a. Sustainability of Data Center Digital Twins with Reinforcement Learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(21): 23832–23834.
- Sarkar, S.; Naug, A.; Luna, R.; Guillen, A.; Gundecha, V.; Ghorbanpour, S.; Mousavi, S.; Markovikj, D.; and Ramesh Babu, A. 2024b. Carbon Footprint Reduction for Sustainable Data Centers in Real-Time. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(20): 22322–22330.
- Sarkar, S.; Naug, A.; Luna Gutierrez, R.; Guillen, A.; Gundecha, V.; Ramesh Babu, A.; and Bash, C. 2023. Real-time Carbon Footprint Minimization in Sustainable Data Centers with Reinforcement Learning. In *NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning*.

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*.