Carbon-Aware LLM Control for Energy-Efficient Liquid-Cooled HPC Data Centers

Avisek Naug[1], **Sahand Ghorbanpour**[1], **Ashwin Ramesh Babu**[1], **Antonio Guillen-Perez**[1], **Vineet Gundecha**[1], **Ricardo Luna Gutierrez**[1], **Soumyendu Sarkar**[1]
[1] HPE Labs, Hewlett Packard Enterprise
{avisek.naug, sahand.ghorbanpour, ashwin.ramesh-babu,antonio.guillen, vineet.gundecha, rluna, soumyendu.sarkar}@hpe.com

Abstract

The rapid growth of large language models (LLMs) in high-performance computing (HPC) data centers necessitates a shift from purely energy-efficient to carbon-aware control for liquid cooling systems. We introduce a novel multi-agent framework that leverages LLM-powered agents to achieve autonomous, carbon-aware thermal management. Our architecture features eight specialized agents coordinated via a hybrid Redis and Model Control Protocol (MCP) backbone for real-time operation. We validate our approach on a high-fidelity digital twin of the Frontier supercomputer's cooling system, focusing on a core contribution: a hybrid Reinforcement Learning (RL) and LLM control strategy. Experimental results show that our 'RL $\rightarrow$ LLM' hybrid model significantly outperforms traditional baselines and other LLM configurations, achieving the lowest average blade temperatures (28.29°C) and the lowest carbon emissions (11.1 kg/hr), while maintaining operational stability. This work presents a practical blueprint for deploying agentic AI to create sustainable, efficient, and explainable control systems for complex cyber-physical infrastructure.

1 Introduction

The growth of AI workloads and large language models (LLMs) in data centers creates significant thermal and energy challenges, necessitating liquid cooling as power densities rise [1, 2, 3, 4]. Optimizing for energy efficiency alone, however, overlooks the substantial carbon footprint. A critical shift towards carbon-aware, multi-objective optimization is required, balancing energy, thermal performance, and sustainability by scheduling workloads to align with renewable energy and addressing the total life-cycle carbon impact [5, 6, 7, 8, 9, 10]. To address this, we propose a control framework using agentic AI systems for proactive management [11, 12, 13, 14]. By integrating multi-agent reinforcement learning (RL) with LLMs, our approach overcomes the poor adaptability and lack of explainability in current systems for carbon reduction [15, 16]. Our agentic architecture uses specialized LLM-agents to manage liquid cooling systems, facilitating scalable, carbon-aware optimization and explainable decision-making through natural language, outperforming conventional controllers in reducing the environmental impact of HPC data centers [17, 18, 19].

2 LLM Driven Holistic Control Architecture

We adopt a multi-agent approach that leverages Large Language Model (LLM)-powered agents to introduce autonomy into our liquid cooling system. The architecture integrates a Redis-based message bus for sub-millisecond inter-agent communication with the Model Control Protocol (MCP) for structured mathematical operations. This hybrid design enables real-time thermal management,

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: .

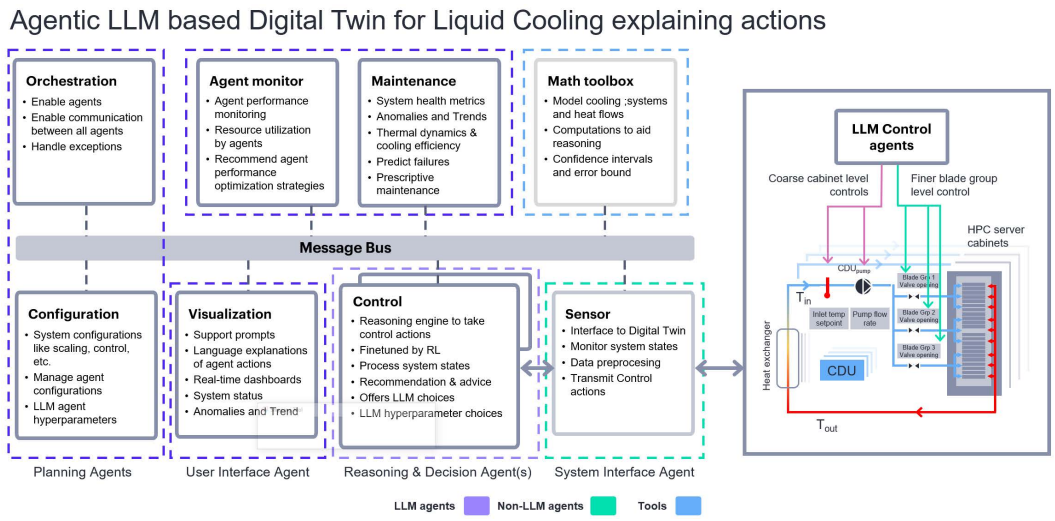

Figure 1: AI Agent implementation of the Liquid Cooling Digital Twin. **Video of the Agentic AI prototype implementation: https://tinyurl.com/llm-agentic-demo**

ensuring system safety and stability while improving operational energy efficiency.
The framework consists of eight specialized agents, including a Control Agent that employs LLM reasoning to generate thermal control actions and a Maintenance Agent that monitors equipment. While the system provides a comprehensive blueprint, our experiments validate the core logic of the Control Agent. By testing controller configurations on a high-fidelity digital twin of the Frontier supercomputer, we demonstrate the effectiveness of our primary contribution: the hybrid RL-LLM strategy serving as the agent's intelligent core.

2.1 Multi-Agent Design

The framework consists of eight specialized agents 1 grouped into categories. **Control agents** manage thermal stability and reliability: the *SensorAgent* processes raw data, the *ControlAgent* uses LLM reasoning for thermal actions, and the *MaintenanceAgent* predicts failures. **Coordination agents** ensure consistent operation: the *OrchestrationAgent* resolves conflicts, the *AgentMonitor* tracks performance, and the *ConfigurationAgent* manages system parameters. Finally, **support agents** enhance usability: the *VisualizationAgent* provides analytics and the *MathToolboxAgent* executes validated computations.

2.2 Hybrid Communication Backbone

The communication architecture combines Redis and MCP, each optimized for distinct roles. Redis facilitates agent-to-agent messaging with sub-millisecond latency, publish–subscribe design, and throughput exceeding 100,000 messages per second. This layer is critical for high-frequency monitoring, emergency stop propagation, and thermal alerts. MCP, by contrast, standardizes tool calls for LLM integration, providing type-safe mathematical computations and schema validation. The hybrid design enables Redis to handle fast coordination while MCP supports structured reasoning tasks such as thermal modeling and predictive maintenance.

2.3 System Description

We demonstrate our method on the open-source Modelica model of the Frontier Supercomputer [20], a framework simulating the thermo-fluidic dynamics of a liquid-cooled exascale data center. The system, illustrated in Figure 2 presents several coupled control problems which we target with our approach.

The first problem domain is the **CDU-Rack Loop**, which consists of 25 pairs of server cabinets and Cooling Distribution Units (CDUs). The control objective here is twofold: 1) to manage the coolant

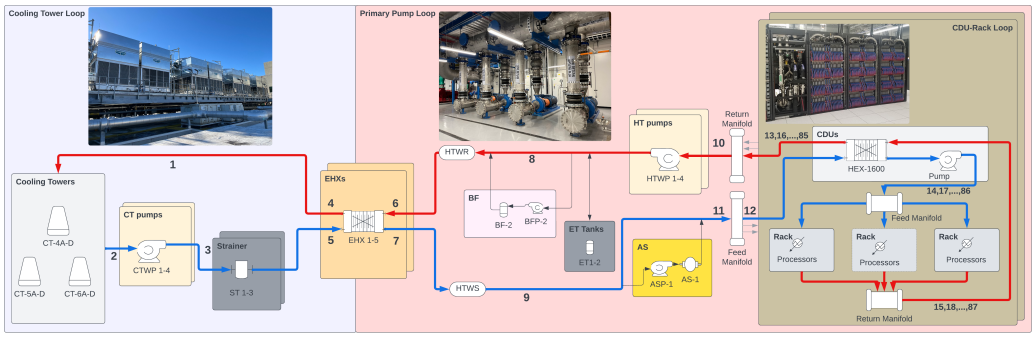

Figure 2: *Frontier's* Cooling System [20]

flow rate and temperature within each CDU, and 2) to actuate manifold valves that distribute coolant to server racks according to their thermal load. Heat from the CDUs is transferred to a central system and ultimately rejected by the **Cooling Tower (CT) Loop**. The second control problem is to optimize the CT's leaving water temperature setpoint. This setpoint critically determines the data center's total cooling capacity and the CT fan's energy consumption, which is influenced by the ambient wet-bulb temperature and the incoming thermal load. To apply LLM inspired control, we wrap the entire simulation model in a Gymnasium environment, enabling control over both the CDU-cabinet and cooling tower subsystems.

3 LLM Integration and Specialization

3.1 Control Agent: Intelligent Thermal Management

At the core of the system lies the **ControlAgent**, which employs LLM reasoning to generate safe and efficient thermal actions. Sensor inputs are translated into structured outputs specifying cabinet-level setpoints, valve positions, and cooling tower actuation. Safety-critical rules are enforced: when blade temperatures fall below $30°$C, the system biases toward positive temperature adjustments and reduced fan operation; when temperatures exceed this threshold, more aggressive cooling responses are triggered. Near-target states are corrected through fine-grained adjustments on the order of $\pm 0.1°$C.

To improve robustness, the agent supports multiple prompting strategies, including chain-of-thought reasoning for stepwise thermal analysis, few-shot learning for optimal policy recall, extended reasoning for complex dynamics, and agent-specific custom prompts.

3.2 Predictive Maintenance

The **SystemHealthAgent** quantifies equipment reliability by analyzing sensor-derived thermal patterns. A composite health score is defined as: $\text{Health Score} = 1.0 - \alpha P_{\text{failure}} - \beta \frac{T_{\text{maint}}}{T_{\text{max}}} - \gamma S_{\text{temp}}$

where P_{failure} denotes the predicted failure probability, T_{maint} the elapsed time since last maintenance, T_{max} the maximum recommended maintenance interval, and S_{temp} a stability measure of thermal fluctuations. This formulation enables proactive scheduling of maintenance tasks while minimizing downtime.

3.3 Real-Time Operation and Live Configuration

The system supports zero-downtime live configuration updates, which are safety-validated, atomically distributed via Redis, and monitored with safety rollbacks. Configurable elements include models, inference parameters, update intervals, reasoning strategies, and routing. Integrated monitoring tracks key performance metrics: thermal stability ($\pm 2°$C), emergency response latency (< 5 ms), uptime ($> 99\%$), energy efficiency gains (10%), and reduced maintenance downtime (20%).

4 Results

4.1 Experiment 1: LLM explanation. guidance and validity

Observation and action tuples from the trained RL policy for the CDU-Cabinet Blade Group are formatted into a structured prompt. For cabinet 1, the observation tuple includes temperatures (T_{1j}) and heat generation rates (P_{1j}) for each of the three blade groups ($j = 1, 2, 3$). The corresponding action tuple contains the valve actuations ($v_{1,j}$), coolant temperature setpoint ($T_{1,cdu}$), and pump differential pressure ($Q_{1,cdu}$). Overall, we notice that the LLM can provide a limited but relevant explanation and guidance for most of the actions, and the system expert generally agrees with the intuitions.

LLM Explainability and Guidance Example

Obs: (307.2, 306.74, 309.77, 21567.55, 44779.22, 33672.41), Act: (0.19, 0.61, 0.2, 22.39, 25.52)
Explanation: Branch 2 remains relatively cool, allowing the controller to redistribute cooling capacity toward more loaded or hotter branches. A setpoint of 22.39 and a controlled flow of 25.52 signify this adaptive behavior. The decision not only maximizes performance-per-watt but also aligns with global decarbonization efforts. It is recommended to reduce branch 1 coolant allocation to 0.1 and to allocate the rest to branch 3 since it has a higher workload
Validity: 1) Human Expert Evaluation: "This is one of the better examples where the agent is explaining the actions for all three blade groups. It however, provides a good way of reifining the RL valve actions"

4.2 Experiment 2: Hybrid RL-LLM Ablation Study

As detailed in Table 1, the **RL → LLM** strategy demonstrates the best performance, achieving the best operational score in the 20–40^{o}C range with lower blade temperature and minimal power draw. While other models like **LLM (Qwen FT + Few-shot)** also performed strongly, the **LLM Base (LLaMA) + Few-shot** was notably ineffective, failing to operate within the target temperature range.

Table 1: Performance comparison of different **LLM and LLM-RL hybrid control strategies**. Lower blade temperature and cooling tower power, along with a higher percentage of temperature steps between 20–40^{o}C, indicate better performance. Experiment uses N=1 tower, m=2 cells per tower, C=5 cabinets with B=3 blade groups per cabinet

Method	**Avg. Blade Temp (C)**	**Cooling Tower Power (W)**	**% Steps 20–40^oC**	**$CO_2(kg/hr)$**
ASHRAE (Baseline)	32.26	26731.31	76.92	12.7
RL	30.65	24131.52	93.28	11.46
LLM Base (LLaMA) + Few-shot	31.42	26500.00	91.25	12.59
LLM Base (Qwen) + Few-shot	30.99	25100.00	93.75	11.92
LLM (LLaMA Fine Tuned by RL traces)	30.20	24001.14	95.33	11.4
LLM (Qwen Fine Tuned by RL traces)	30.01	23378.22	95.97	**11.1**
LLM (LLaMA FT + Few-shot)	29.72	24689.92	95.12	11.73
LLM (Qwen FT + Few-shot)	29.37	23750.74	96.08	11.28
LLM (Qwen FT) → RL Hybrid	30.31	26759.77	93.97	12.71
RL → LLM (Qwen FT) Hybrid	**28.29**	**23371.92**	**96.80**	**11.1**

5 Conclusion

This paper introduced a carbon-aware, multi-agent control system for liquid-cooled HPC data centers. Our experiments, conducted on a high-fidelity digital twin of the Frontier exascale system, demonstrated that our 'RL → LLM' approach is highly effective. It successfully minimized carbon footprint and blade temperatures while ensuring operational stability. The system's architecture, built on a fast and reliable communication backbone, proves the feasibility of using agentic AI for complex, real-time decision-making and provides a valuable mechanism for generating human-readable explanations for control actions. Our findings provide a robust blueprint for developing the next generation of autonomous, sustainable, and transparent control systems for critical infrastructure.

References

[1] Qinghao Hu, Zhisheng Ye, Zerui Wang, Guoteng Wang, Meng Zhang, Qiaoling Chen, Peng Sun, Dahua Lin, Xiaolin Wang, Yingwei Luo, Yonggang Wen, and Tianwei Zhang. Characterization of large language model development in the datacenter. *NSDI*, 2024.

[2] Interplex. How the rise of generative ai is impacting data centers and network infrastructures. *Interplex Trends*, 2024.

[3] Flex Power Modules. The basics of liquid cooling in ai data centers. *Flex Power Modules Blog*, 2024.

[4] Mikros Technologies. Ai and ml liquid cooling solutions. *Mikros Technologies*, 2024.

[5] Yang Zhao, Zhelun Chen, and Liang Wang. Efficient multi-agent reinforcement learning hvac power consumption optimization. *SSRN Electronic Journal*, 2024.

[6] Daniel Bayer and Marco Pruckner. Enhancing the performance of multi-agent reinforcement learning for controlling hvac systems. *arXiv preprint arXiv:2309.06940*, 2023.

[7] Soumyendu Sarkar, Avisek Naug, Ricardo Luna Gutierrez, Antonio Guillen, Vineet Gundecha, A Ramesh Babu, and Cullen Bash. Real-time carbon footprint minimization in sustainable data centers with reinforcement learning. In *NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning*, 2023.

[8] Abrar Hossain, Abubeker Abdurahman, Mohammad A Islam, and Kishwar Ahmed. Power-aware scheduling for multi-center hpc electricity cost optimization, 2025.

[9] Hayden Moore, Sirui Qi, Ninad Hogade, Dejan Milojicic, Cullen Bash, and Sudeep Pasricha. Sustainable carbon-aware and water-efficient llm scheduling in geo-distributed cloud datacenters, 2025.

[10] Yueying Li, Zhanqiu Hu, Esha Choukse, Rodrigo Fonseca, Suh G Edward, and Udit Gupta. Ecoserve: Designing carbon-aware ai inference systems, 2025.

[11] McKinsey. Seizing the agentic ai advantage, 2025.

[12] IBM. What is agentic ai? *IBM Think*, 2025.

[13] StackGen. Meet the 7 ai agents that build, govern, heal, and optimize infrastructure. *StackGen Blog*, 2025.

[14] Van Jones. Building the future: Advances in ai infrastructure for autonomous agents. *Wellington Management*, 2025.

[15] Liang Zhang and Zhelun Chen. Large language model-based interpretable machine learning control in building energy systems. *Energy and Buildings*, 2024.

[16] BrainBox AI. How llms are revolutionizing building management. *BrainBox AI Blog*, 2024.

[17] Michael Knight. Agentic ai: The self-healing datacenter and the critical human roles. *LinkedIn*, 2025.

[18] 75F. Irem: How conversational ai transforms building management. *75F News*, 2024.

[19] Ben Bartling. Can the llm be used to setup building automation? *LinkedIn*, 2024.

[20] Wesley Brewer, Matthias Maiterth, Vineet Kumar, Rafal Wojda, Sedrick Bouknight, Jesse Hines, Woong Shin, Scott Greenwood, David Grant, Wesley Williams, and Feiyi Wang. A digital twin framework for liquid-cooled supercomputers as demonstrated at exascale. In *SC24: International Conference for High Performance Computing, Networking, Storage and Analysis*, page 1–18. IEEE, November 2024.

A Technical Appendices and Supplementary Material