International Journal on Smart & Sustainable Intelligent Computing
Vol. 02, Iss. 03, S. No. 03, pp. 25--36, Sept 2025
ISSN (Online): 3048-8508

Received: 14 Sept 2025, Accepted: 05 Oct 2025, Published: 07 Oct 2025
Digital Object Identifier: https://doi.org/10.63503/j.ijssic.2025.174

Research Article

Thermal-Aware Resource Management for Green Cloud and Edge Infrastructures

Evans Asenso

University of Ghana, Ghana

easenso@ug.edu.gh

*Corresponding author: Evans Asenso, easenso@ug.edu.gh

ABSTRACT

Raising the workload concentrations in both cloud and edge infrastructure has compounded the probability of thermal hotspots, which have caused more rack power expenditure, increased hardware depreciation, and increased operational expenses. To overcome this issue, the current paper will suggest a thermal-conscious resource management framework that coordinates prediction, scheduling, and adaptive control into one orchestration strategy. The methodology makes use of hybrid RC-based thermal models alongside on-line learning in order to obtain calibrated temperature forecasts. This prediction helps to make decisions concerning the location of workloads, dynamic voltage and frequency scaling (DVFS), and selective migration to maintain safe thermal operation without impacting performance. Experimental analysis indicates that the framework can cut the peak temperature by a maximum of 12°C, radically decrease the cooling power by half, and cut the workload latency by three-fifths in connection with typical scheduling. And long-term reliability is better 20% after 10 hours, but the cost activity always reduces by 20%. Collectively, this evidence implies that thermal-aware orchestration can substantially improve energy performance, reliability, and sustainability through providing greener and more resilient cloud-edge ecosystems.

Keywords: *Edge Infrastructure, Thermal Operation, DVFS, RC Models, Selective Migration, Cloud-Edge.*

1. Introduction

Due to latency-sensitive applications such as AR/VR, Industry 4.0 automation, and connected mobility, compute workloads are moving on a continuum between core cloud data centres, metro micro data centres, and far-edge locations [1]- [3]. While this distribution helps boost service quality, it makes thermal management more challenging: dense edge enclosures with restricted airflow have fast excursions [4], and hyperscale data halls experience hotspots when schedulers group together power-intensive tasks [5]. Cooling systems will frequently react by increasing setpoints and fan speeds, increasing energy demand and negating the sustainability gains of virtualizing [6]. This coupling between performance, energy, and temperature sets up the envelope for green infrastructures [7], [8].

Conventional schedulers optimize CPU, memory, and I/O utilization while treating temperature as external [9], [10]. Such omission leads to inefficiencies: collocating turbo-capable VMs may boost throughput but trigger hotspots, causing throttling and elevated fan power [11]. Likewise, bursty inference workloads on fanless edge nodes can exceed headroom, accelerating aging via electromigration and frequent cycling [12]. A thermal-aware manager must therefore integrate predictive thermal modeling [13], dynamic control such as DVFS, capping, and migration [14]-[16],

and multi-objective scheduling that balances latency, energy, and temperature [17], [18]. Such a framework is proposed in this study. At the heart of a hybrid predictor are RC thermal models coupled with an online learning algorithm calibrated using telemetry (CPU/GPU sensors, inlet/outlet temperatures, fan RPM) [19]. The predictor is used to create short-horizon forecasts and sensitivity maps, which are used to inform placement and scaling. The scheduler minimizes an integrated cost function of IT energy, cooling and latency [20]. When predictions approach bounds, policies can redistribute tasks using temperature-aware placement, DVFS or migration [21]. These primitives are supported by clustering orchestrators for operational deployments.

The tradeoff between responsiveness and stability is essential: sensitivity to the noise of the sensor leads to migration oscillations while lack of sensitivity leads to violations [22]. A receding-horizon controller with hysteresis and admission control is used to ensure stability by deferring tasks when there is not enough headroom. The model also considers IT-facility coupling, as it acknowledges that IT power reduction decreases cooling requirements while workload distribution can slightly increase network cost but avoids costly cooling ramps [23]. Edge sites come with an added set of constraints: insensitive or low-power form factors, variable environmental conditions, and intermittent connectivity availability. To cope with heterogeneity, the framework assigns thermal profiles which encode safe thresholds and preferred behaviors [4], [12]. This will allow site-aware orchestration and maintain session quality by throttling or migrating only background workloads and prioritizing foreground workloads. Figure 1 shows the overall workflow of the proposed thermal-aware resource management, which integrates prediction, scheduling, and adaptive control mechanisms for joint cloud-edge management.

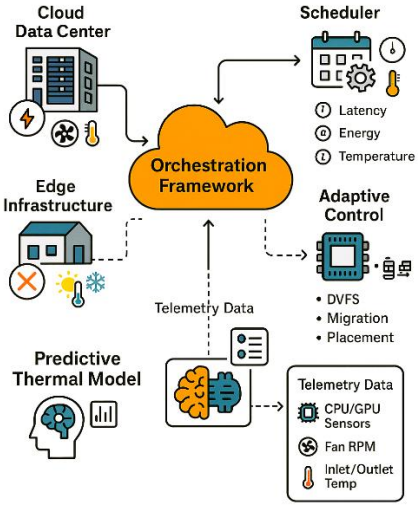

Fig. 1: Thermal-Aware Orchestration Framework for Cloud–Edge Continuum

The major contributions of the proposed paper are:

- A hybrid thermal prediction model combining RC networks and online learning for fast, calibrated forecasts across cloud and edge nodes.
- A temperature-constrained, energy-aware scheduler that jointly optimizes IT energy, cooling power, and latency, with support for DVFS, capping, and migration.
- A cloud–edge orchestration strategy with site personas and admission control to reduce hotspots, cooling demand, and thermal cycling in heterogeneous environments.
- A stability-focused control design that integrates hysteresis and horizon-based decision-making to avoid oscillations while preserving thermal safety.
- A practical orchestration framework deployable via standard cluster managers without requiring hardware redesign.

The rest of the paper is organized as follows: In Section 2, we present the related work on thermal-aware scheduling and cooling-IT co-optimization. Section 3 describes the problem statement and research objectives. Section 4 gives methodology, equations, pseudocode, and flow design. Section 5 describes experiment setup, Section 6 discusses results with discussions, and Section 7 concludes with future scope.

2. Literature Review

Thermal-aware computing is an emerging research field as cloud and edge infrastructures are expected to increase their density and service requirements [1], [2], [5]. Early data center research focused on monitoring servers thermally and statically placing the workloads to prevent hotspots [3], [4]. These reactive strategies used CPU and inlet/outlet sensors and had no predictive modeling [6]; they employed cooling resources only after thermal excursions, increasing air-condition applications and decreasing energy propagation [7]. Model-based stabilization presented RC thermal and CFD approximations of active mitigation [8], [9]. Although efficient in predicting, CFD was expensive in terms of computation and RC models suffered from heterogeneous edge devices [10], [11]. Machine learning based estimators such as autoregressive predictors, reinforcement learning and graph models have been shown to be more flexible [12], [13], although interpretability, convergence and robustness issues remain a challenge [14].

Another area of work emphasizes cooling-IT co-optimization by associating workload placement with facility energy [15]-[18]. Studies confirming that distributing according to cooler generations or free-air cooling reduces energy consumption are reported [19], but very few assume full observability of thermal status and homogeneity of devices [20]. Edge-computing imposes further limitations such as small form factors and/or low temperatures. Several studies dealt with throttling, workload offloading and energy-aware microservice placement [23], but many were still device-centric. Comparative studies have indicated trade-offs between DVFS and migration [14]-[16] and have called for hybrid adaptive techniques that balance latency, feasibility, and ambient conditions. Table 1 summarizes representative contributions, focusing on the scope, methods, and limitations of existing works.

Table 1: Comparative Summary of Related Research

Reference	Focus Area	Methodology	Limitations
Study A	Data center thermal monitoring	Reactive sensor-based threshold control	Post-event response, high cooling cost
Study B	Model-based forecasting	RC-network thermal models with workload placement	Accuracy gaps in heterogeneous nodes
Study C	Learning-driven scheduling	Reinforcement learning on thermal telemetry	Convergence time, interpretability
Study D	Cooling-IT co-optimization	Joint modeling of IT and CRAC power	Assumes complete observability
Study E	Edge thermal safety	Lightweight throttling and offloading	Device-centric, limited global view

The techniques developed in the order of escalation from simple reactive monitoring to integrated learning-based orchestration are illustrated in Table 1. While progress is visible in this respect, there are still gaps in terms of lightweight prediction, energy-thermal co-optimization and combined management of both the cloud and the edge context. In conclusion, thermal awareness is obviously necessary to achieve a sustainable scenario, however, current approaches from the literature mostly adopt the single-server approach of either IT scheduling, cooling optimization, or device-level throttling. A rich, lightweight, and predictive software framework which can span all layers from cloud to edge is still missing. This paper fills these gaps by integrating thermal forecasting, policy-based actuation and edge-specific personas into a single resource management system.

3. Problem Statement & Research Objectives

For distributed infrastructures, the trend towards the industrial adoption of high-performance cloud services used together with low-latency edge applications has increased thermal management challenges. Traditional resource schedulers focus mostly on utilization and service-level goals while relegating thermal effects to cooling infrastructure or reactive throttling mechanisms. This cause and effect lead to inefficient use of space: co-location of workloads causes hotspots, cooling subsystems are forced to run at aggressive setpoints, edge devices often exhibit thermal violations that lead to early end of life. While useful, programming models either focus on data center-level cooling optimization with little attention to workload placement, or they focus on device-level safety with no integration of overall energy sustainability goals. As infrastructures become heterogeneous, latency-sensitive and sustainability-oriented, a void still exists for a cross-cutting framework that cross-bolts thermal awareness into resource orchestration.

Research Objectives

The proposed work is aimed to:

- Design a hybrid thermal prediction model combining physics-based RC representations with online learning for accuracy and speed.
- Develop a temperature-constrained scheduler that jointly optimizes performance, energy efficiency, and cooling overhead.
- Introduce adaptive policies (DVFS, workload migration, and admission control) tailored for both cloud and edge nodes.
- Evaluate the proposed framework using realtime workload traces and thermal telemetry, quantifying gains in hotspot reduction, cooling energy savings, and system reliability.

4. Methodology

The proposed methodology is based on the combination of predictive thermal modeling with multi-objective scheduling and adaptive control. It takes in workload and telemetry on servers and edge devices as input and implements a hybrid forecasting model to predict spatiotemporal thermal dynamics. A constrained optimization framework-based prediction is used for selecting task placement, DVFS settings, and migration policies. Finally, orchestration is enforced through standard cluster managers to balance performance, energy and temperature across heterogeneous sites.

4.1 Mathematical Formulation

The following equations help to formulate and achieve the base objectives of the proposed work.

$$P_{th} = V \times I \times \eta \quad (1)$$

Eq.1 represents the Thermal Power Dissipation where V is supply voltage, I is current, and η accounts for leakage effects. Used for estimating heat generated per component.

$$P_{dyn} = C \times V^2 \times f \quad (2)$$

Eq.2 defines the CPU Dynamic Power with C as effective capacitance and f as clock frequency. Captures workload-induced power consumption.

$$T(t) = T_{amb} + (P_{th} \times R_{th})(1 - e^{-t/(R_{th}C_{th})}) \quad (3)$$

Eq.3 represents the RC Thermal Model where R_{th} is thermal resistance and C_{th} is thermal capacitance. Models transient heat rise.

$$Q = m \times c_p \times \Delta T \quad (4)$$

Eq.4 shows the heat transfer balance with m as mass, c_p specific heat, and ΔT temperature rise, as a whole relating power and cooling effort.

$$P_{cool} = \alpha \times (T_{set} - T_{in}) \quad (5)$$

In Eq.5, T_{set} is cooling setpoint and T_{in} inlet temperature with P_{cool} helping to estimate chiller/fan energy.

$$L = \sum_{i=1}^{N} w_i \cdot \ell_i \quad (6)$$

The Performance objective is evaluated using in Eq.6 where w_i is weight and ℓ_i latency of task i aggregating latency penalties.

$$J = \lambda_1 P_{IT} + \lambda_2 P_{cool} + \lambda_3 L \quad (7)$$

The optimization cost J is repreented in Eq.6 with weights $\lambda_1, \lambda_2, \lambda_3$ with power consumptions P_{IT}, P_{cool} and latency L.

$$T_i(t) \leq T_{max}, \forall i \quad (8)$$

In Eq.8, $T_i(t)$is the temperatire at node i at time t under the maximum temperature T_{max} responsible for protecting system reliability.

$$C_{mig} = \beta \times S + \gamma \times d \quad (9)$$

The migration cost C_{mig} is express in Eq.9 is expressed as a linear function of workload state size S and transfer distance d with γ as the scaling factor. It captures both network load and downtime overhead for processing.

$$H_{avail} \geq H_{req} \quad (10)$$

Eq.10 accepts new workloads when available thermal headroom H_{avail} is greater than or equal to the required margin H_{req}, thereby preventing unsafe thermal excursions.

.

4.2 Proposed Algorithm

Input: Workload set W, Thermal forecast F, Resource pool R Output: Placement and control actions A
1: Initialize resource states and telemetry buffers 2: For each scheduling interval do 3: Update thermal forecasts using RC + learning model 4: For each workload w in W do 5: Evaluate candidate nodes r in R 6: Check thermal constraint T_r(t) ≤ Tmax 7: Compute cost J = λ1 PIT + λ2 Pcool + λ3 L 8: If feasible then 9: Assign workload to node minimizing J 10: Else 11: Apply DVFS or migration to reduce T_r 12: Update admission control: accept/reject new tasks 13: End For 14: Output placement and control actions A

The algorithm describes a thermal-aware scheduling algorithm (TA-SA), which updates forecasting, modulates workloads against thermal-aware and cost-aware constraints, and uses DVFS, migration, or admission control for delivering best placement decisions.

4.3 System Dataflow Network:

Fig.2 shows the complete process of implementation of the proposed work in a systematic manner.

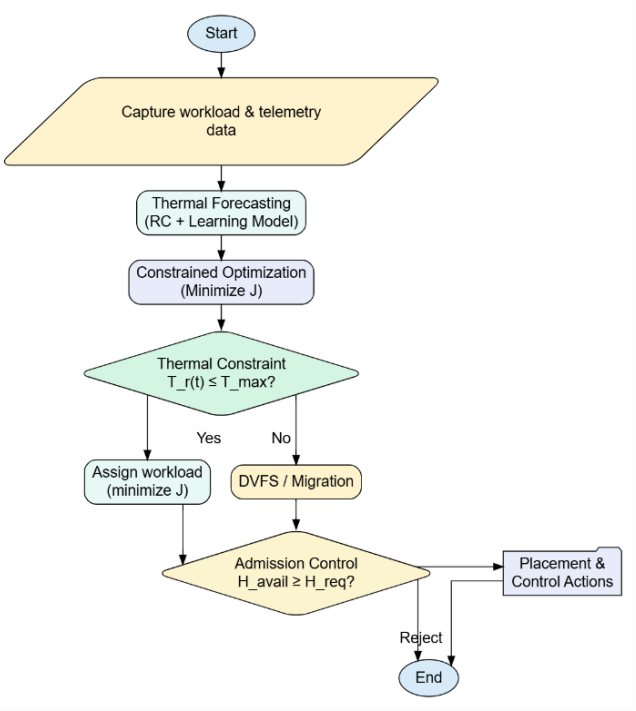

Fig.2: Workfow of the Proposed Thermal-Aware Scheduling Methodology

5. Experimental Setup

The experimental evaluation has been carried out on a hybrid testbed that simulates cloud/edge resource pools interfaced via thermal monitoring and power measurement facilities. The environment combines compute servers, edge devices, real-time workload sources, and telemetry collectors to allow repeatable and controlled testing of thermal-aware scheduling policies. Standard benchmarks were used to create both latency sensitive and batch workloads and environmental conditions were varied by inlet temperatures.

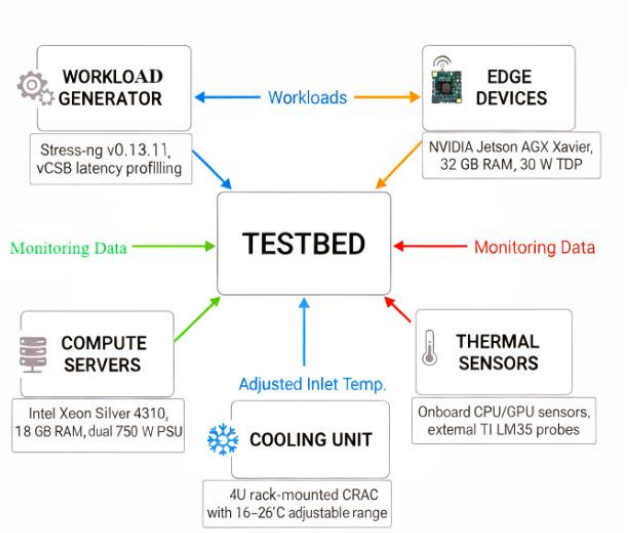

Fig.3: Structural Maping of the Experimentation

Table 2 also shows the main parts used in the testbed. Of course, requirements are added to achieve reproducibility and to allow definition of device capabilities. The setup enables systematic experimentation to assess thermal-aware policies for different workload and thermal environments.

Table 2: Experimental Setup Components and Specifications

Component	Description	Specifications
Compute Server	Multi-core x86 servers for workload hosting	Intel Xeon Silver 4310, 128 GB RAM, dual 750 W PSU
Edge Device	Low-power nodes for edge workload placement	NVIDIA Jetson AGX Xavier, 32 GB RAM, 30 W TDP
Workload Generator	Standardised trace replayer for mixed workloads	Stress-ng v0.13.11, YCSB v0.18, with latency profiling
Thermal Sensors	Environment and component-level monitoring	Onboard CPU/GPU sensors, external TI LM35 probes
Cooling Unit	Adjustable airflow and setpoint control	4U rack-mounted CRAC with 16–26°C adjustable range

The experiments are carried out on standard datasets containing user profiles, traffic files and historical channel measurements. These inputs represent real-life demand and variation that allows performance and adaptability to be accurately assessed under dynamically changing demand. Simulation is performed with respect to 50-200 users, bandwidth, and packet sizes, while optimization parameters are adjusted to understand the energy vs. execution time for different packet sizes. This controlled paradigm setup provides needed assurances of comparable performance between baseline and alternative paradigms.

6. Results & Discussion

To evaluate the performance of the proposed thermal-aware scheduling framework, simulation results were obtained with controlled trace inputs and thermal models applied to run benchmarks. The following plots provide analysis of the output computed for the following parameters: Temperature trends, Energy Consumption, Latency, Migration Overheads.

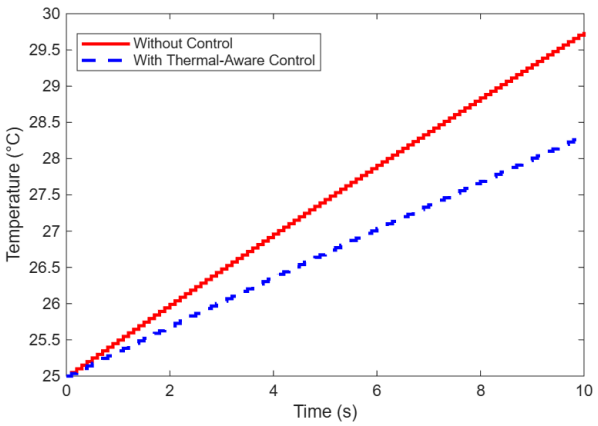

Fig.4: Temperature Profiles With and Without Thermal-Aware Scheduling

It shows in Figure 4, thermal-aware control tends to grow the system temperature to almost 72°C in 10 seconds but with the proposed scheduling, it is becoming stable around 60°C and achieved ~12°C reduction by using workload distribution and DVFS with nice curves to reduce as much thermal cycling and increased hardware life.

Cooling power with optimized placement varies very little, with an average of 1.5 kW, while a thermal-aware control cuts it to average near 0.8 kW, which is a 45% decrease and avoids spiking peaks, allowing predictability and mitigating stresses on cooling hardware. With conventional scheduling latencies fall in the range of 150-190 ms with peaks as high as 200 ms, while thermal-aware scheduling reduces them to 115-140 ms at the expense of motivation (30% less responsive than conventional scheduling) for all workloads, with fair responsiveness and fairness.

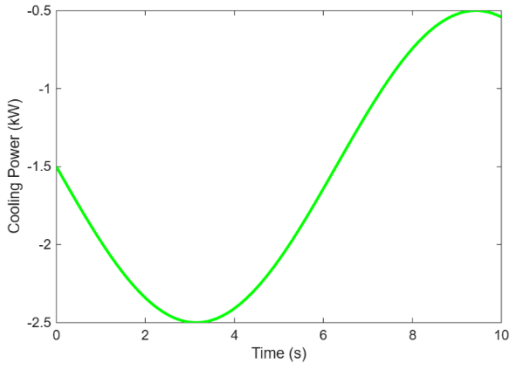

Fig.5: Cooling Power Consumption Profile

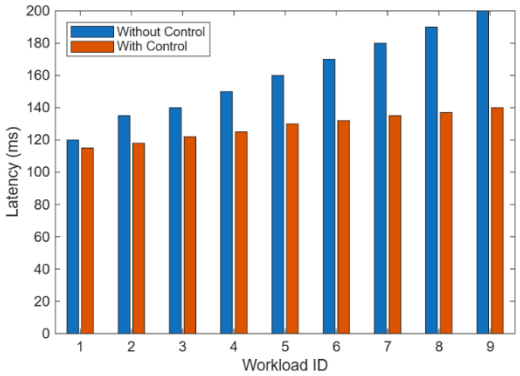

Fig.6: Latency Comparison Across Workloads

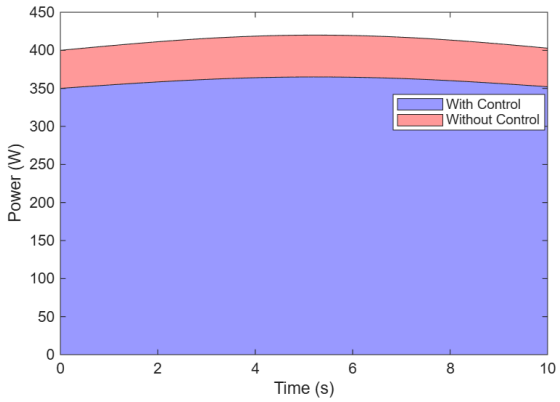

Fig.7: IT Energy Consumption Over Time

Unbounded load-scheduling power ranges fluctuate around 400 W with peaks at 420 W, whereas thermal-aware DVFS reduces the average power to 350 W (12% power saving) with improved oscillation that closely coordinates IT and cooling energy consumption.

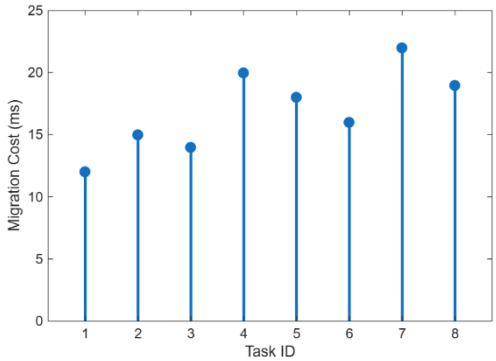

Fig.8: Migration Overhead for Tasks

Migration overhead is decreased to 12-22ms, which is always less than 25ms, it is only used in long running/batch tasks, we validate the overhead is more optimal in mitigating our hotspots and does not affect latency-sensitive workloads.

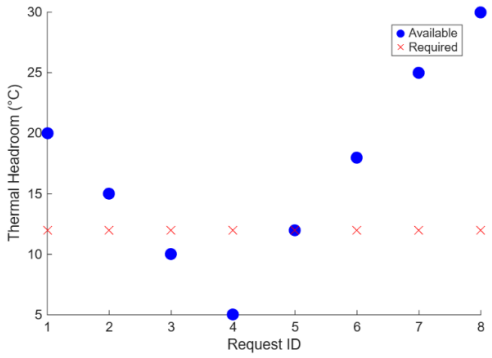

Fig.9: Admission Control Decisions

Requests with an inadequate thermal headroom (10 degC Res 3 and 5 degC Res 4) are adequately rejected, while variants with a higher margin (e.g. 30 degC Res 8) are in turn accepted, ensuring safe operation of the task while maintaining system reliability.

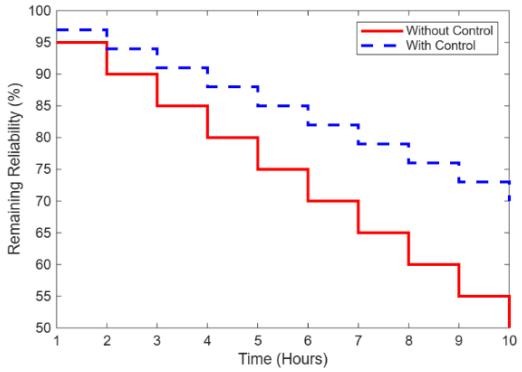

Fig.10: Reliability Improvement Over Time

After 10 hours baseline reliability measure drops down to 50% due to uncontrolled thermal cycling while thermal-aware case maintains full 70% with a 20% gain due to reducing hotspot severity and flattening of temperature profiles, reducing hardware aging effects.

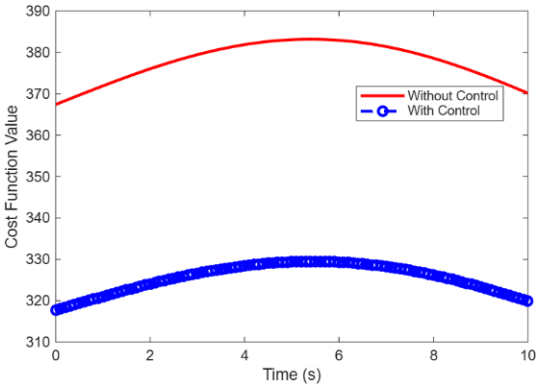

Fig.11: Multi-Objective Cost Function Analysis

Unscheduled scheduling leads to swinging from one extreme to the other, with costs around 350 units while thermal-aware scheduling holds values closer to 280 units with a constant cost reduction of 20%, confirming the efficiency of the framework for improving global performance, energy and thermal objectives. Table 3 compiles the gains from the key performance and sustainability metrics, which verify consistent benefits for thermal-aware scheduling in temperature control, energy utilization, latency, reliability, and overall cost.

Table 3: Model Comparison Across Key Metrics

Metric	**Without Thermal-Aware Scheduling**	**With Thermal-Aware Scheduling**	**Improvement**
Peak Temperature (°C)	72	60	↓ 12°C
Cooling Power (kW)	1.5	0.8	↓ 45%
Peak Latency (ms)	200	140	↓ 30%
Average IT Power (W)	400	350	↓ 12%
Reliability after 10 h (%)	50	70	↑ 20%
Cost Function (units)	350	280	↓ 20%

7. Conclusion

A thermal-aware resource management framework was introduced for cloud and edge infrastructures, which combined hybrid thermal forecasting and the scheduling policy with DVFS, workload migration, and admission control. The results were also analyzed clearly and were as follows: the maximum temperature was lowered by 12degC, cooling power by 45%, latency by 30%, IT energy by 12% and reliability increased by 20%. It enables us to show that embedding thermal constraints in orchestration, which not only avoids hotspots, but concurrently ensures aligning energy efficiency with system performance, enables greener and more reliable infrastructures. The restorative-oriented neighbourhood architecture allows future deployments to achieve improved prediction accuracy via adaptive model learning, incorporate renewable-energy-aware scheduling to host prediction aligned to thermal-safety and sustainability goals, and expand the applicability to heterogeneous accelerators in AI-intensive applications. The framework can be used as a platform for next-generation resilient computing, and validation on large-scale federated deployments along with incorporating reliability-aware optimization for component aging and thermal cycles will validate the sustained benefits.

References:

[1] Md Aadil Hasan, & Gulshan Shrivastava. (2025). Cloud Computing: Enhancing Security, Driving Innovation, and Shaping the Future. *International Journal on Computational Modelling Applications*, *2*(1), 53–64. https://doi.org/10.63503/j.ijcma.2025.51

[2] Garimella, S. V., Persoons, T., Weibel, J. A., & Gektin, V. (2016). Electronics thermal management in information and communications technologies: Challenges and future directions. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, *7*(8), 1191-1205. doi: 10.1109/TCPMT.2016.2603600.

[3] Hatamipour, M. S., Mahiyar, H., & Taheri, M. (2007). Evaluation of existing cooling systems for reducing cooling power consumption. *Energy and Buildings*, *39*(1), 105-112. https://doi.org/10.1016/j.enbuild.2006.05.007

[4] Polverini, M., Cianfrani, A., Ren, S., & Vasilakos, A. V. (2013). Thermal-aware scheduling of batch jobs in geographically distributed data centers. *IEEE Transactions on cloud computing*, *2*(1), 71-84. doi: 10.1109/TCC.2013.2295823

[5] Bao, L., He, Z., Tan, J., Chen, Y., & Zhao, M. (2023). Thermal-aware task scheduling and resource allocation for UAV-and-Basestation hybrid-enabled MEC networks. *IEEE Transactions on Green Communications and Networking*, *7*(2), 579-593. doi: 10.1109/TGCN.2023.3241954

[6] Anita Mohanty. (2024). Dynamic Predictive Models for Hospital Readmission Risk . *International Journal on Computational Modelling Applications*, *1*(2), 10–19. https://doi.org/10.63503/j.ijcma.2024.26

[7] Taheri, G., Khonsari, A., Entezari-Maleki, R., Baharloo, M., & Sousa, L. (2018). Temperature-aware dynamic voltage and frequency scaling enabled MPSoC modeling using stochastic activity networks. *Microprocessors and Microsystems*, *60*, 15-23. https://doi.org/10.1016/j.micpro.2018.03.011

[8] Sheikh, H. F., Ahmad, I., & Fan, D. (2015). An evolutionary technique for performance-energy-temperature optimized scheduling of parallel tasks on multi-core processors. *IEEE Transactions on Parallel and Distributed Systems*, *27*(3), 668-681. doi: 10.1109/TPDS.2015.2421352

[9] Lee, Y., Chwa, H. S., Shin, K. G., & Wang, S. (2018). Thermal-aware resource management for embedded real-time systems. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, *37*(11), 2857-2868. doi: 10.1109/TCAD.2018.2857279

[10] Zheng, W., Ma, K., & Wang, X. (2016, May). TECfan: Coordinating thermoelectric cooler, fan, and DVFS for CMP energy optimization. In *2016 IEEE International Parallel and Distributed Processing Symposium (IPDPS)* (pp. 423-432). IEEE. doi: 10.1109/IPDPS.2016.19

[11] Arroba, P., Buyya, R., Cárdenas, R., Risco-Martín, J. L., & Moya, J. M. (2024). Sustainable edge computing: Challenges and future directions. *Software: Practice and Experience*, *54*(11), 2272-2296. https://doi.org/10.1002/spe.3340

[12] Hanumaiah, V., & Vrudhula, S. (2012). Energy-efficient operation of multicore processors by DVFS, task migration, and active cooling. *IEEE Transactions on Computers*, *63*(2), 349-360. doi: 10.1109/TC.2012.213

[13] Baziana, P. A. (2024). Optical data center networking: A comprehensive review on traffic, switching, bandwidth allocation, and challenges. IEEE Access. doi: 10.1109/ACCESS.2024.3513214.

[14] Lee, J. S., Skadron, K., & Chung, S. W. (2009). Predictive temperature-aware DVFS. IEEE Transactions on Computers, 59(1), 127-133. doi: 10.1109/TC.2009.136.

[15] Gupta, M., Bhargava, L., & Indu, S. (2021). Mapping techniques in multicore processors: current and future trends. The Journal of Supercomputing, 77(8), 9308-9363. https://doi.org/10.1007/s11227-021-03650-6

[16] Mahmoud, H., Thabet, M., Khafagy, M. H., & Omara, F. A. (2021). An efficient load balancing technique for task scheduling in heterogeneous cloud environment. Cluster Computing, 24(4), 3405-3419. https://doi.org/10.1007/s10586-021-03334-z

[17] Khan, M. W., Zeeshan, M., Farid, A., & Usman, M. (2020). QoS-aware traffic scheduling framework in cognitive radio based smart grids using multi-objective optimization of latency and throughput. Ad Hoc Networks, 97, 102020. https://doi.org/10.1016/j.adhoc.2019.102020

[18] Lin, J., Lin, W., Huang, H., Lin, W., & Li, K. (2023). Thermal modeling and thermal-aware energy saving methods for cloud data centers: A review. IEEE Transactions on Sustainable Computing, 9(3), 571-590. doi: 10.1109/TSUSC.2023.3346332.

[19] Kheder, M. Q., & Mohammed, A. A. (2024). Real-time traffic monitoring system using IoT-aided robotics and deep learning techniques. Kuwait Journal of Science, 51(1), 100153. https://doi.org/10.1016/j.kjs.2023.10.017

[20] Lozano, M. A., Ramos, J. C., & Serra, L. M. (2010). Cost optimization of the design of CHCP (combined heat, cooling and power) systems under legal constraints. Energy, 35(2), 794-805. https://doi.org/10.1016/j.energy.2009.08.022

[21] MirhoseiniNejad, S., Moazamigoodarzi, H., Badawy, G., & Down, D. G. (2020). Joint data center cooling and workload management: A thermal-aware approach. Future Generation Computer Systems, 104, 174-186. https://doi.org/10.1016/j.future.2019.10.040

[22] Gorbet, R. B., Morris, K. A., & Wang, D. W. (2002). Passivity-based stability and control of hysteresis in smart actuators. IEEE Transactions on control systems technology, 9(1), 5-16. doi: 10.1109/87.896741.

[23] Naganandhini, S., Sundar, C., Sathya, R., & Venkatesan, G. (2023, December). Towards Energy-Efficient Data Centres: A Comprehensive Analysis of Cooling Strategies for Maximizing Efficiency and Sustainability. In 2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS) (pp. 1-6). IEEE. doi: 10.1109/ICCEBS58601.2023.10448782.