Increasing Energy Efficiency of Server Cooling Over Traditional Methods with Deep Reinforcement Learning Agents Running on OCP Compliant BMC Platforms

AI-ML model for dynamic server fans speed control achieves better energy efficiency than the traditional fans control methods. Model runs on an ML engine of a BMC chip.

Scaling Innovation Through Collaboration

OCP REGIONAL SUMMIT
24-25 APRIL 2024
LISBON, PORTUGAL

Increasing Energy Efficiency of Server Cooling Over Traditional Methods with Deep Reinforcement Learning Agents running on OCP Compliant BMC platforms

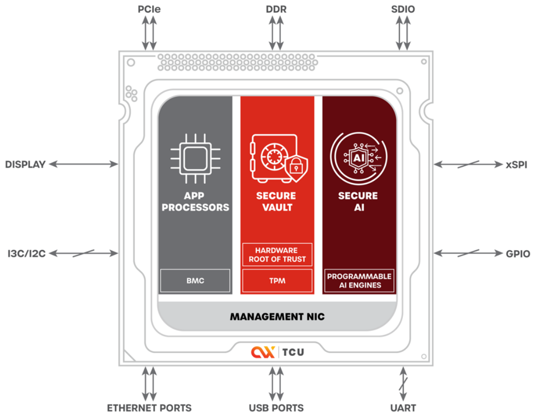
SUSTAINABILITY

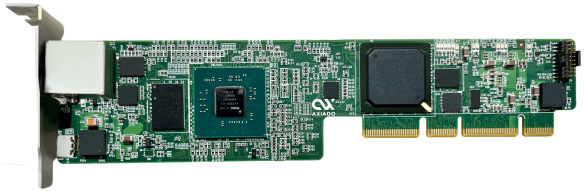
Raghu Kondapalli, Chief Technology Officer, Axiado Corporation
Sundaram Arumugasundaram, Principal Security Architect, Axiado Corporation
Zhichao Zhang, Principal Machine Learning Architect, Axiado Corporation

OPEN COMMUNITY®

Trusted Control/Compute Unit (TCU) — Overview

- TCU chip consists of below components:
  1. App processors: cores for running apps like BMC, host vulnerability management, extended detection and Response (XDR) agents
  2. programmable AI engine to run ML models like server thermal management
  3. Smart-NIC for control/management plane like BMC traffic
  4. hardware Root of trust(HRoT) and TPM(Trusted Platform Module) to enhance server security
- Axiado offers Smart-SCM that is compliant with the Open Compute Project (OCP) datacenter-ready secure control module (DC-SCM) standard.





Next-Gen Thermal Management: The Power of ML on TCU/BMC — Overview

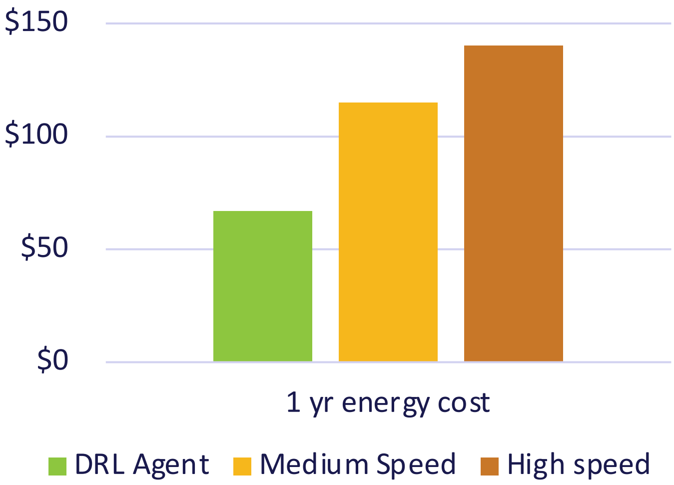
- **AI-Powered Dynamic Thermal Management (DTM) from BMC:** BMC is ideal for server thermal management due to its existing role in various server management functions, including power control.
- **Faster Thermal Prediction and Calibration:** TCU collects sensor data directly, bypassing the host OS, enabling faster thermal prediction and fan speed calibration.
- **Rich Dataset for Decision Making:** As an OCP DC-SCM compliant BMC, TCU gathers comprehensive data from all chassis components (CPUs, GPUs, etc.) via diverse connections (I2C, eSPI, USB, PCI-e), providing a rich dataset for optimal fan control decisions.

Next-Gen Thermal Management: The Power of ML on TCU/BMC — Overview

- **Dedicated ML engine for DTM-ML model:** As TCU's ML engine only runs the DTM-ML model, it offers timely inference and fan speed control
- **Hardware-Based Security:** Leveraging confidential computing and other security features, TCU protects the DTM-ML model from potential vulnerabilities on the host OS, offering a more secure solution.
- **Proactive Management with PMC Data:** TCU utilizes CPU and GPU Performance Monitoring Counters (PMC) to proactively manage thermals based on workload demands.
- Integration of AXIADO's DTM-ML with the open standards like openBMC and DMTF's redfish is work in progress.

ML-DTM Result – Significant Energy Savings

Fan Energy Savings: Energy savings up to **50%**
Annual savings per server: **$70**
Annual savings for 100K servers: **$7 million**



	FAN cycles	Kilowatt-hour per hour	1 Year ($0.10 per kwh)
TCU DTM-ML	Optimized	0.076	$67
Medium Fan Speed	65.5%	0.131	$115
High Fan Speed	80%	0.160	$140

DTM-ML model training and deployment Details

- **Data Collection:** pulling data from sensors every 5 seconds for six months
- **Trained on random and diverse intensity workloads** with a massive data set
- **Analysis and Prediction ML type:** DRL (Deep Reinforcement Learning)
- **Continuous Learning:** improves energy efficiency over time.
- **Surpasses PID Fan Controllers:** delivers superior results to PID controllers through broader dataset correlation. Unlike reactive PID controllers, it proactively adjusts fan speeds based on anticipated workload demands.

Benefits of DRL(Deep Reinforcement Learning)

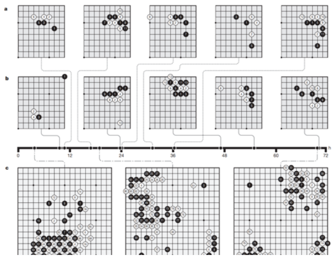
- DRL is a revolutionary AI methodology that combines reinforcement learning and deep neural networks. By iteratively interacting with an environment and making choices that maximize cumulative rewards, it enables agents to learn sophisticated strategies, directly learn rules from sensory inputs, which makes use of deep learning's ability to extract complex features from unstructured data https://www.geeksforgeeks.org/what-is-reinforcement-learning/.
- DTM-DRL self learns from the environment, continuously improves the efficiency of the balance between temperature and energy usage and proactively anticipates cooling needs based on workloads and other dynamic factors.

Reinforcement Learning – solving many complex problems

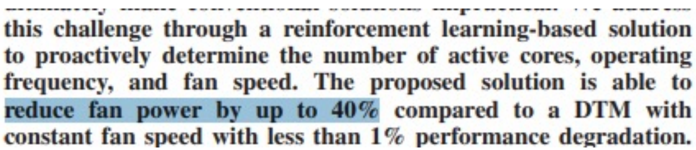
AlphaGo Zero

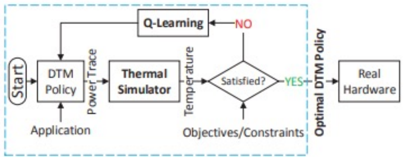
Mastering the game of Go without human knowledge

Reinforcement learning to learn the meaning of states from the environment



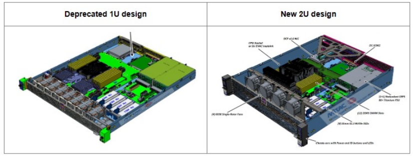
Dynamic Thermal Management with Proactive Fan Speed Control Through Reinforcement Learning


this challenge through a reinforcement learning-based solution to proactively determine the number of active cores, operating frequency, and fan speed. The proposed solution is able to reduce fan power by up to 40% compared to a DTM with constant fan speed with less than 1% performance degradation.



One rule can't fit them all

To get the best saving, policies need to be learnt from different environment, hardware and workload



AI for Google data center cooling – success story

- https://blog.google/inside-google/infrastructure/safety-first-ai-autonomous-data-center-cooling-and-industrial-control/

Rules don’t get better over time, but AI does.

AI-Powered Dynamic Thermal Management (DTM)

- Reinforcement Learning redefines DTM by replacing heuristic human input with self-optimizing AI agents.
  - **Human vs. AI Agency:** From manually tuned protocols to AI-driven, Q-Learning-based autonomous agents.
  - **AI Superiority:** RL agents predictive management cuts fan power by 40%.
- Outcome: Autonomous agents offer continual learning, precision, and efficiency, redefining DTM in data-centric environments.

Diverse Policies Learned by Deep Reinforcement Learning

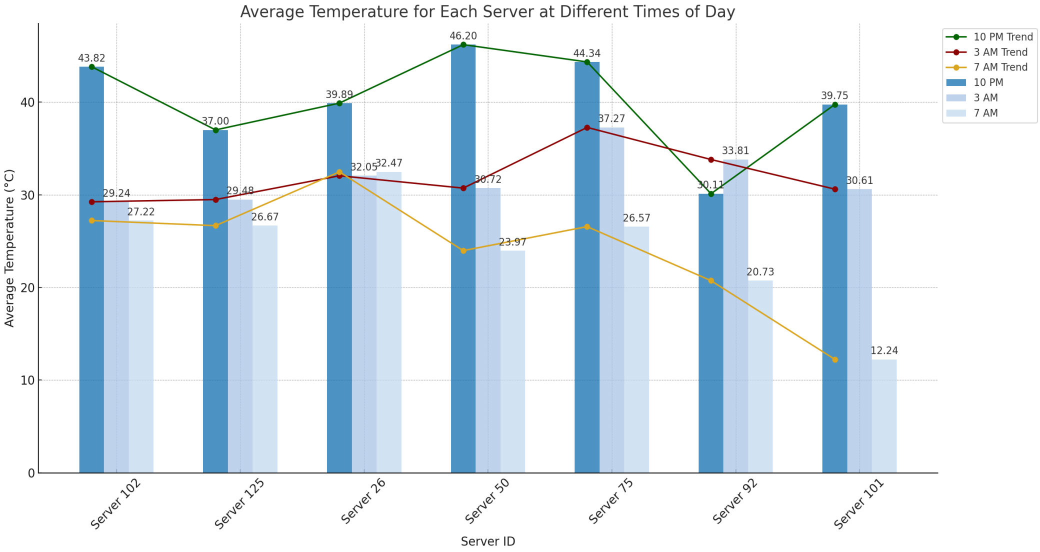
Temperature

Fan Speed

1. **Precision Control:** The RL model develops fine-tuned cooling algorithms, directly improving energy management.
2. **Intelligent Adaptation:** It swiftly adjusts to fluctuations, ensuring consistent performance under varying load conditions.
3. **Sustainable Operations:** Forecasts and adjusts to future demands, significantly reducing the carbon footprint and operating costs.

Server Room Temperature Monitoring

- The real-time temperature data manually collected from Axiado HQ's server room, it's evident that server temperatures vary significantly not only by rack position but also by time of day. For instance, the consistent decrease in temperatures from 10 PM to 7 AM across most servers suggests ambient factors, possibly related to lower night-time room temperatures or reduced server activity, greatly influence server temperatures.
- Leveraging this data can inform efficient cooling, power usage, cost reduction and other server optimization strategies.



Summary — Redefining Data Centers with AI-Driven DTM

- **TCU BMC Integration:** A power-efficient BMC controller that is OCP compliant, equipped with an on-chip NPU, requiring only 0.5TOPs for this application. This integration is not just a step towards modernization but a leap towards cost-effective and green computing, given the necessity of BMC controllers in modern data centers
- **Smart Scaling:** Tailored AI dynamically adapts to diverse server configurations, ensuring optimal performance across any data center layout.
- **Operational Excellence Reimagined:** Transition from traditional, labor-intensive methods to AI-driven strategies. Our real-world deployments demonstrate how integrating AI with real-time sensor data and machine learning not only enhances system reliability but also significantly reduces operational costs.
- **Energy Efficiency & Sustainability:** Leveraging AI for real-time control of cooling systems results in up to 40% savings on cooling energy costs. This approach not only slashes energy bills but also substantially reduces the carbon footprint, contributing to greener data center operations.

AI-Driven DTM PUE Impact

- Achieving up to **18.6% PUE** Improvement
  - Reallocating fan power the PUE changed from 1.09 to 1.61
  - A 50% reduction in fan power leads to a new PUE of 1.31.
    - With a 5% reduction in fan power, the new PUE is 1.57
    - With a 50% reduction in fan power, the new PUE is 1.31
- Air cooling at the server level is widely used in data centers, particularly for configurations up to 10 kW per rack. While alternative cooling methods are gaining traction for higher-density setups, server-level air cooling remains a common, cost-effective choice.

Initial PUE Calculation:

- PUE is defined as the ratio of Total Facility Energy to IT Equipment Energy.

$$\text{PUE} = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}}$$

Given Assumptions:

- Current PUE is 1.09.
- IT Equipment Energy is 80% from servers.
- Server Energy Breakdown:
  - Fan Power: 200W
  - CPU Power: 400W
  - Other Components: 50W
- Total Server Energy:

$$\text{Total Server Energy} = \text{Fan Power} + \text{CPU Power} + \text{Other Power}$$
$$\text{Total Server Energy} = 200W + 400W + 50W = 650W$$

- IT Equipment Energy (80% from servers):

$$\text{IT Equipment Energy} = \frac{\text{Total Server Energy}}{0.80}$$
$$\text{IT Equipment Energy} = \frac{650W}{0.80} = 812.5W$$

Calculating Non IT Facility Energy:

- Solve for Non IT Facility Energy from the PUE formula:

$$\text{Non IT Facility Energy} = (\text{PUE} - 1) \times \text{IT Equipment Energy}$$
$$\text{Non IT Facility Energy} = (1.09 - 1) \times 812.5W$$
$$\text{Non IT Facility Energy} = 0.09 \times 812.5W = 73.125W$$

New PUE Calculation with Fan Power Reallocated:

- Adjusted IT Equipment Energy by subtracting the fan power:

$$\text{Adjusted IT Equipment Energy} = \text{Total Server Energy} - \text{Fan Power}$$
$$\text{Adjusted IT Equipment Energy} = 650W - 200W = 450W$$

- Adjusted Non IT Facility Energy by adding the fan power:

$$\text{Adjusted Non IT Facility Energy} = \text{Non IT Facility Energy} + \text{Fan Power}$$
$$\text{Adjusted Non IT Facility Energy} = 73.125W + 200W = 273.125W$$

- New PUE with the fan power reallocated:

$$\text{New PUE} = 1 + \frac{\text{Adjusted Non IT Facility Energy}}{\text{Adjusted IT Equipment Energy}}$$
$$\text{New PUE} = 1 + \frac{273.125W}{450W} \approx 1.6069$$

Impact of Fan Power Reduction on New PUE:

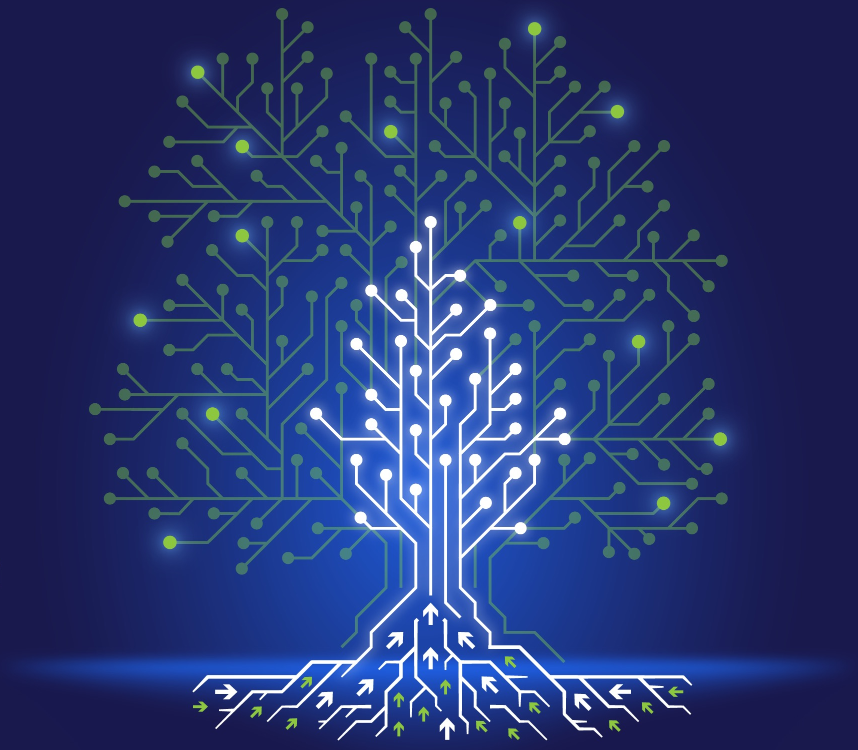
- For each reduction percentage (from 5% to 50% reduction in fan power):

$$\text{Reduced Fan Power} = \text{Fan Power} \times (1 - \text{Reduction Percentage})$$
$$\text{New Adjusted IT Equipment Energy} = \text{Total Server Energy} - \text{Reduced Fan Power}$$
$$\text{New Adjusted Non IT Facility Energy} = \text{Non IT Facility Energy} + \text{Reduced Fan Power}$$
$$\text{New PUE with Fan Power Reduction} = 1 + \frac{\text{New Adjusted Non IT Facility Energy}}{\text{New Adjusted IT Equipment Energy}}$$

Call for Action

- Problem to Solve
  - Let’s collaborate to create an ML based fans speed control as part of OCP, OpenBMC and DMTF and save energy.
- How to get involved in the Project
  - By piloting the deployment of the DTM-ML model in your data center.
- Timeline for Contribution Availability
  - From now to end of 2025
- Timeline for Product Availability
  - From now to end of 2025
- Where to find additional information (URL links):
  - Work In Progress

Thank you!



Scaling Innovation Through Collaboration

OCP REGIONAL SUMMIT

24-25 APRIL 2024
LISBON, PORTUGAL