EE&AE
2025 10th International Conference
on Energy Efficiency and Agricultural Engineering
5-7 November 2025, Starozagorski Bani, Bulgaria
IEEE

Energy-Efficient Task Scheduling in Data Centers using Adaptive Deep Reinforcement Learning

Dinuț Ionuț-Cosmin, Rodica-Claudia Constantinescu and Bogdan Alexandrescu

Faculty of Electronics, Telecommunications and Information Technology
National University of Science and Technology POLITEHNICA Bucharest, Romania
ionut_cosmin.dinut@stud.etti.upb.ro, r.constantinescu@upb.ro,
bogdan.alexandrescu@upb.ro

GOAL OF THE STUDY

Efficiently managing energy consumption while maintaining low service latency is a critical challenge for modern data centers. This paper presents a novel approach using a Deep Reinforcement Learning (RL) agent, trained with Proximal Policy Optimization (PPO), to schedule incoming tasks across multiple servers dynamically.

METHODOLOGY OF THE INVESTIGATION

- **Framework:** The system is modeled as a Markov Decision Process (MDP), where the agent learns to make optimal decisions.
- **Algorithm:** We use Proximal Policy Optimization (PPO) to train the agent in a custom-built, dynamic simulation environment.
- **Core Logic:** The agent learns a policy that balances two competing goals by maximizing a reward function designed to:
  - Minimize energy consumption.
  - Reduce SLA violations.

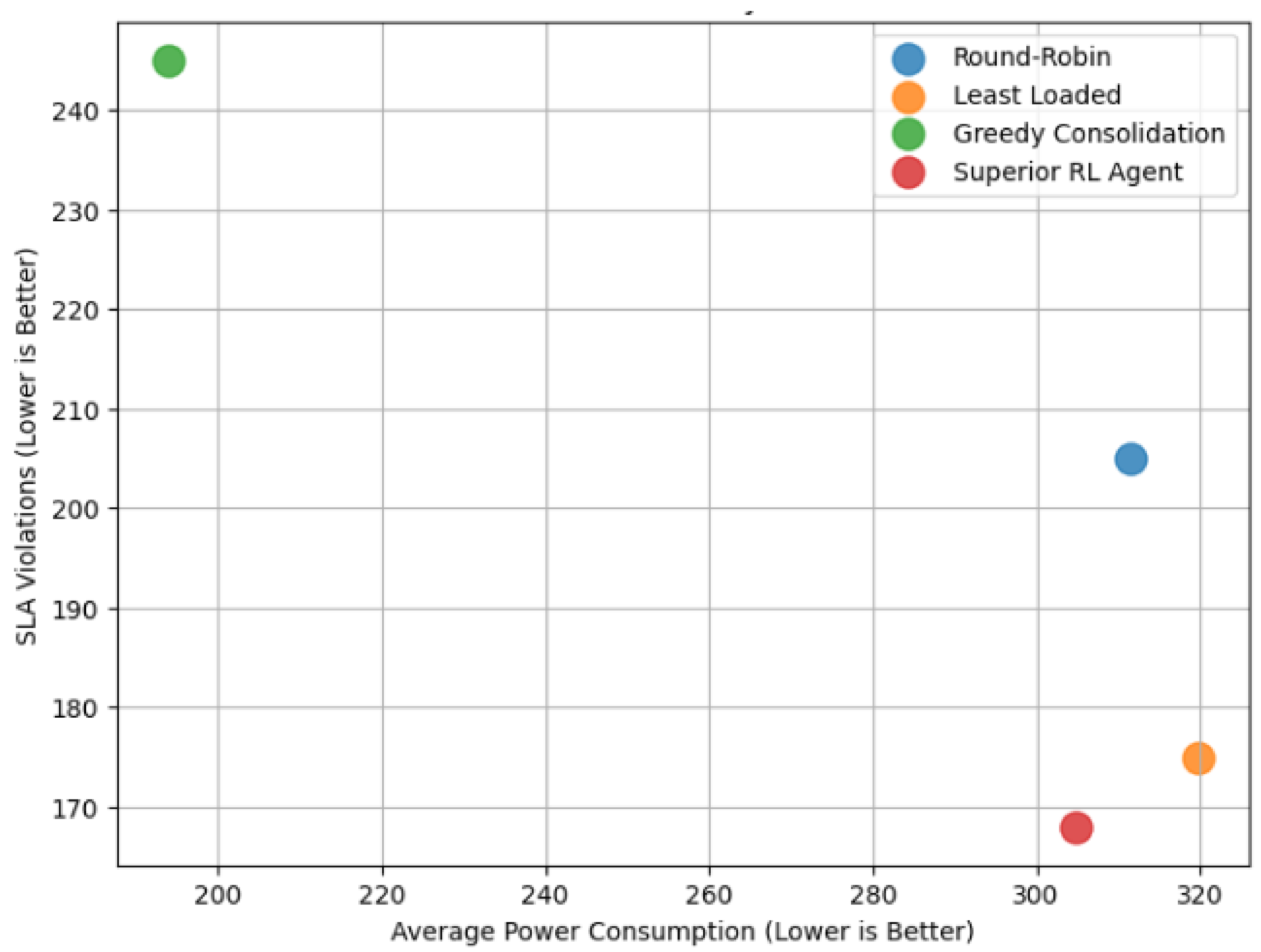

Fig. 1. Aggregate Performance: Power vs Reliability trade-off

Table 1. Comparative Performance Metrics For All Evaluated Policies

Policy	Average Power (W)	Task Completed	SLAs Violated
Greedy Consolidation	193.92	95	245
Round-Robin	311.38	349	205
Least Loaded	319.64	381	175
Superior RL Agent	403.80	369	168

Table 2. Hyperparameters used for PPO training

Hyperparameter	Value
Learning Rate	3e-4
Discount Factor (γ)	0.99
GAE Lambda (λ)	0.95
Batch Size	64
PPO Clip Range (ε)	0.2
Training Timesteps	800

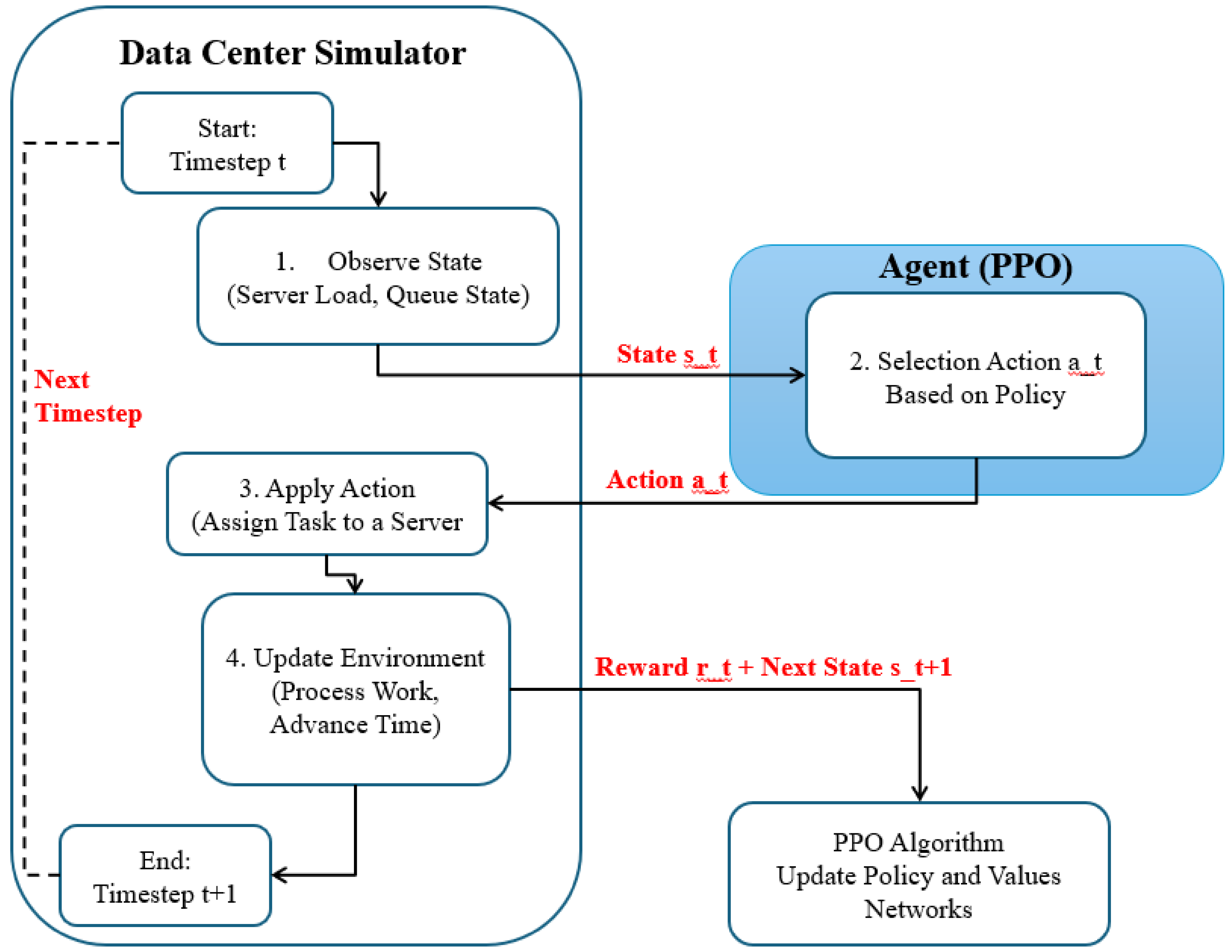

Fig. 2. Block diagram of the agent-environment interaction loop

CONCLUSIONS

Achievement: We successfully trained a Deep Reinforcement Learning agent that consistently outperforms static scheduling policies.

Key Finding: The agent learns a hybrid strategy, which saves energy at low loads and maximizes performance at high loads.

Impact: This work validates RL as a powerful and practical solution for autonomous resource management in data centers.