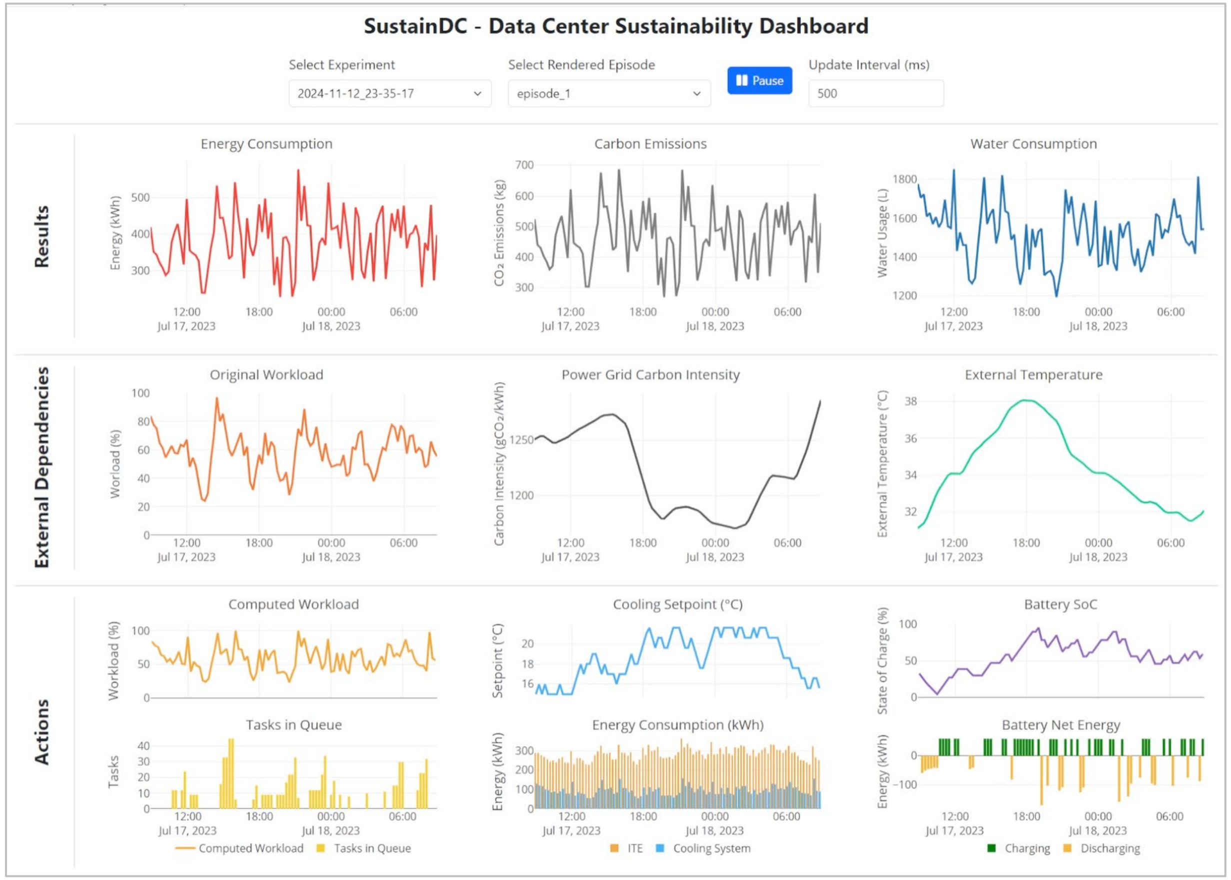
SustainDC: Benchmarking for Sustainable Data Center Control

Avisek Naug · Antonio Guillen-Perez · Ricardo Luna Gutierrez · Vineet Gundecha · Cullen Bash · Sahand Ghorbanpour · Sajad Mousavi · Ashwin Ramesh Babu · Dejan Markovikj · Lekhapriya Dheeraj Kashyap · Desik Rengarajan · Soumyendu Sarkar

Hewlett Packard Labs @ Hewlett Packard Enterprise



GitHub

OpenReview

NeurIPS 2024

email: soumyendu.sarkar@hpe.com

Motivation

NEURAL INFORMATION PROCESSING SYSTEMS

Holistic Real-time Data Center Optimization for Sustainability with increased AI workloads

Sustainability Goals

- Lower Carbon Emissions
- Lower Energy Consumption
- Lower Water Usage

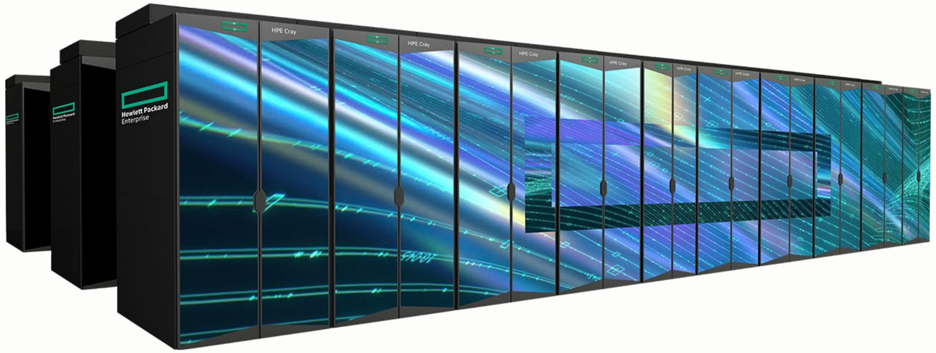
Paradigm Shift in Energy Optimization with Multiple controls

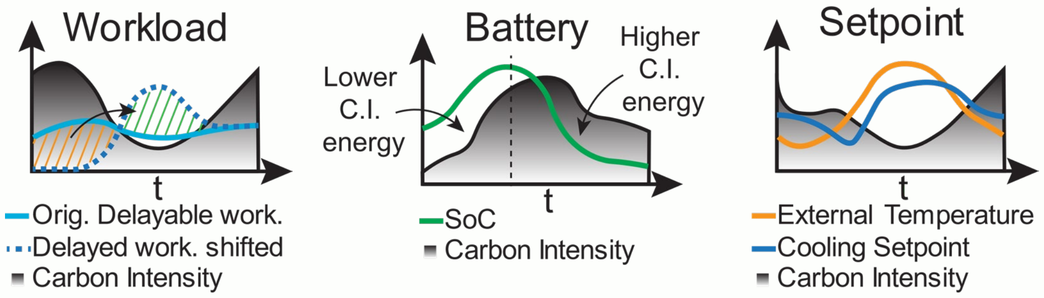
- Optimize Cooling with IT Energy
- Schedule Flexible Loads
- Energy Storage
- Multi-agent Real-time Optimization

Holistic Real-Time Optimization of Data Centers

Resolving Multiple External + Inter-Module Dependencies

- Weather and Grid Carbon Intensity
- Dependency between Cooling, Scheduling, and Energy Storage



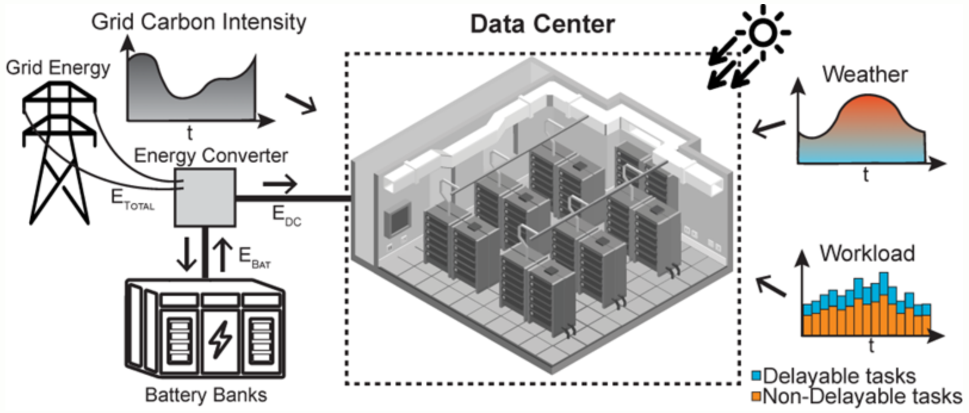


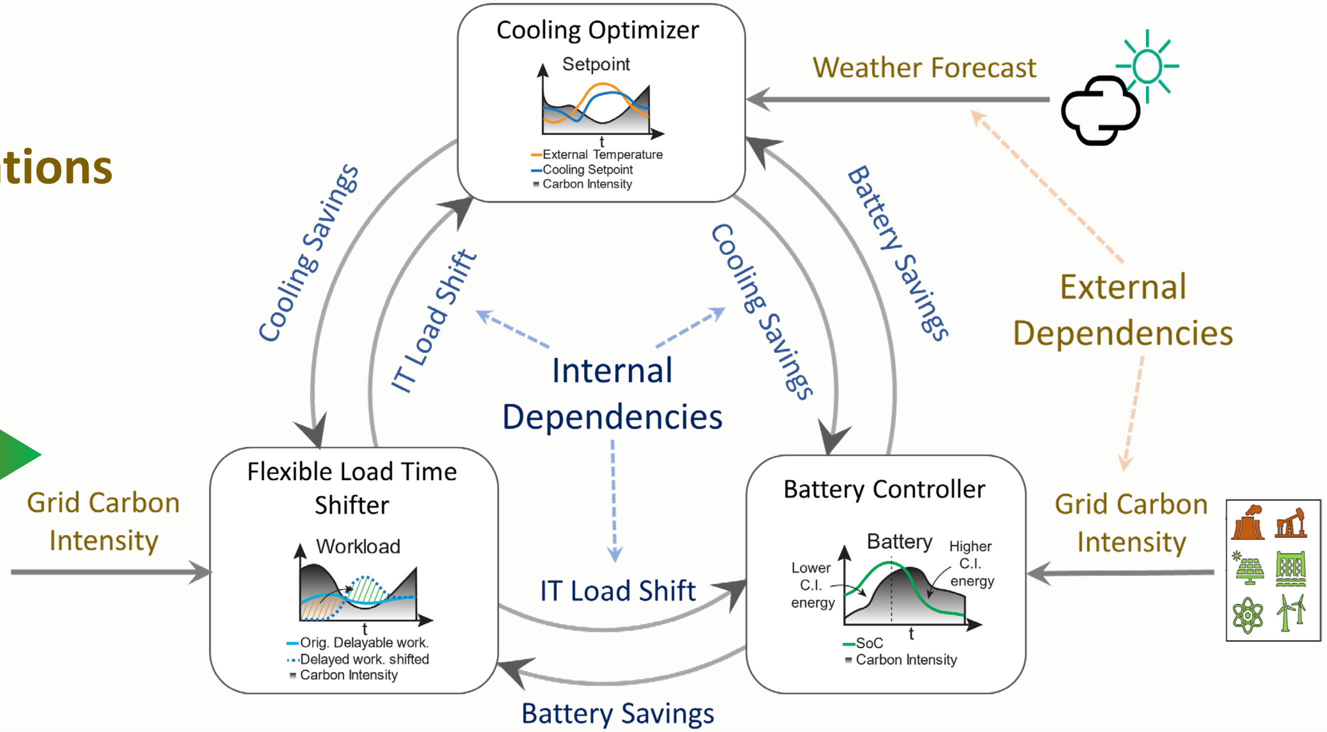
Problem: Control Challenges

NEURAL INFORMATION PROCESSING SYSTEMS

Holistic Optimization requires resolving multiple Internal and External Dependencies:

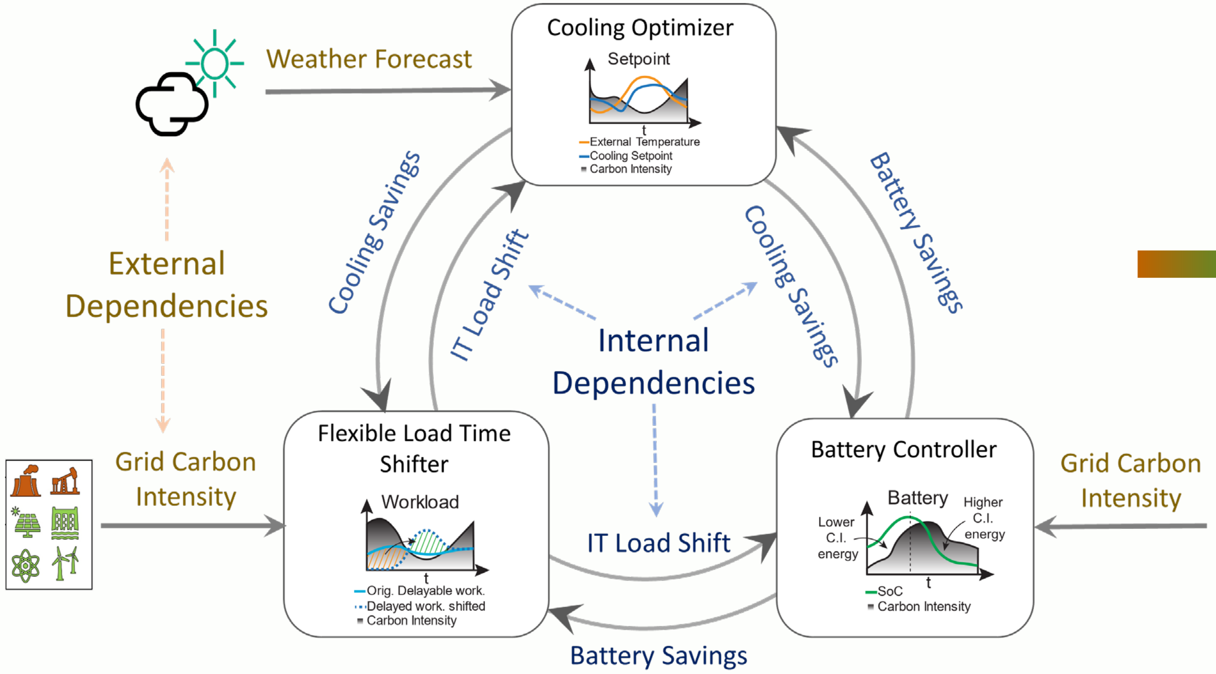
- Internal Dependencies
  - Cooling & IT Energy
  - Dynamic Flexible Workload Scheduling
  - Energy Storage
- External Dependencies
  - Weather and Grid Carbon Intensity variations

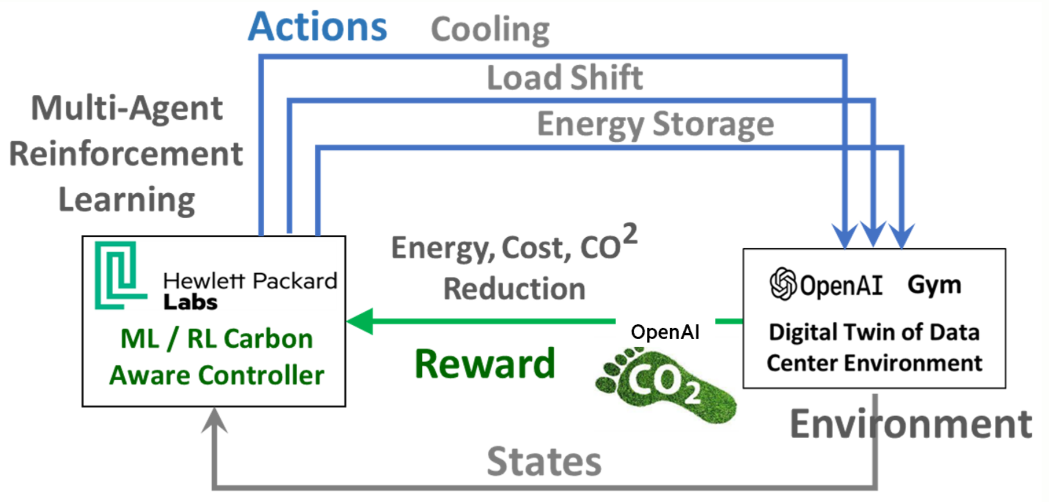




Solution

Holistic Optimization requires Multi-agent Multi-objective Reinforcement Learning Control



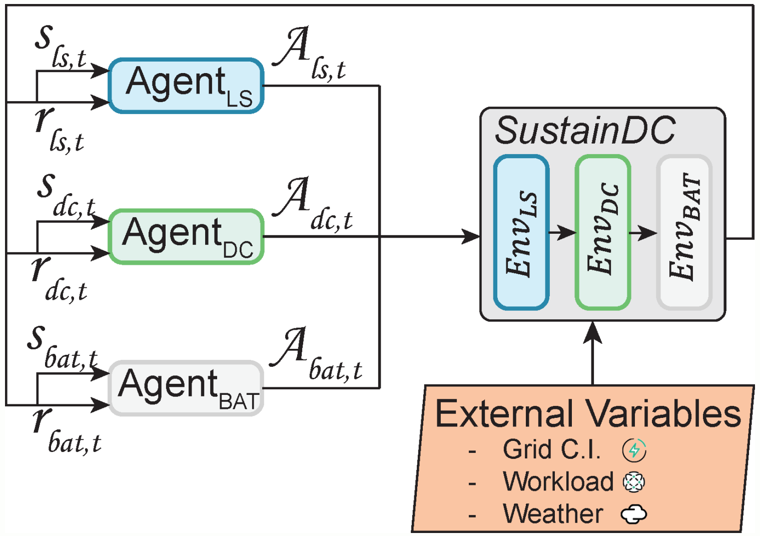


Multi-Agent Reinforcement Learning Controller

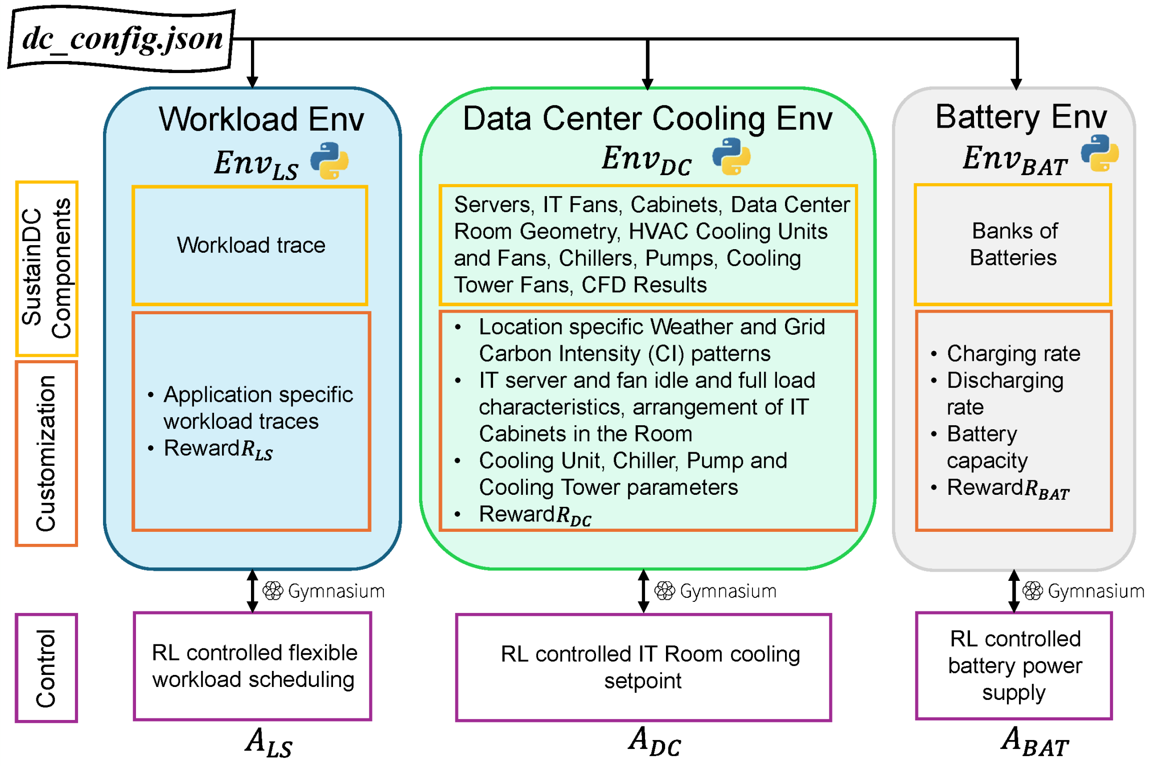
SustainDC: Comprehensive RL Benchmark Environment

Customizable RL Environment for Optimization

- Workload Environment
- Data Center Cooling Environment
- Battery Environment
- Customizable configurations



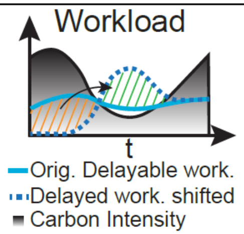
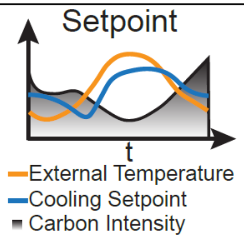
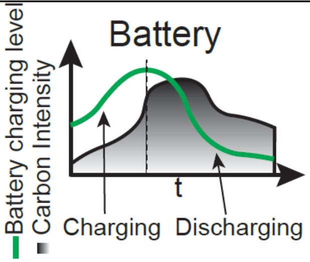
RL agents in SustainDC



SustainDC with the three main environments - Workload Env, Data Center Cooling Env, and Battery Env along with their customizable components and control actions

3 Interacting Control Problems & RL Agents

- **Workload scheduling decisions**
- **Cooling setpoint optimization**
- **Battery charging/discharging strategy**

Agent	Control Knob	Actions	Optimization Strategy	Figure
Agent_{LS}	Delay-tolerant workload scheduling	$\begin{cases} 0 & \text{Store Delayable Tasks} \\ 1 & \text{Compute All Immediate Tasks} \\ 2 & \text{Maximize Throughput} \end{cases}$	Shift tasks to periods of lower CI/lower external temperature/other variables to reduce the CFP.	
Agent_{DC}	Cooling Setpoint	$\begin{cases} 0 & \text{Decrease Setpoint} \\ 1 & \text{Maintain Setpoint} \\ 2 & \text{Increase Setpoint} \end{cases}$	Optimize cooling by adjusting cooling setpoints based on workload, external temperature, and CI.	
Agent_{BAT}	Battery energy supply/store	$\begin{cases} 0 & \text{Charge Battery} \\ 1 & \text{Hold Energy} \\ 2 & \text{Discharge Battery} \end{cases}$	Store energy when CI/temperature/workload/other is low and use stored energy when is high to reduce CFP.	

$$
\begin{aligned}
Agent_{LS} &: (SC_t \times CI_t \times D_t \times B_t) \rightarrow A_{ls,t} \\
Agent_{DC} &: (SC_t \times t_{db} \times t_{room} \times E_{hvac} \times E_{it} \times CI_t) \rightarrow A_{dc,t} \\
Agent_{BAT} &: (SC_t \times Bat\_SoC \times CI_t) \rightarrow A_{bat,t} \\
Env_{LS} &: (B_t \times A_{ls,t}) \rightarrow \hat{B}_t \\
Env_{DC} &: (\hat{B}_t \times t_{db} \times t_{room} \times A_{dc,t}) \rightarrow (E_{hvac}, E_{it}) \\
Env_{BAT} &: (Bat\_SoC \times A_{bat,t}) \rightarrow (Bat\_SoC, E_{bat}) \\
CFP_t &= (E_{hvac} + E_{it} + E_{bat}) \times CI_t
\end{aligned}
$$

$$
\left(\theta_{LS}, \theta_{DC}, \theta_{BAT}\right) = argmin\left(\sum_{t=0}^{t=N} CFP_t\right)
$$

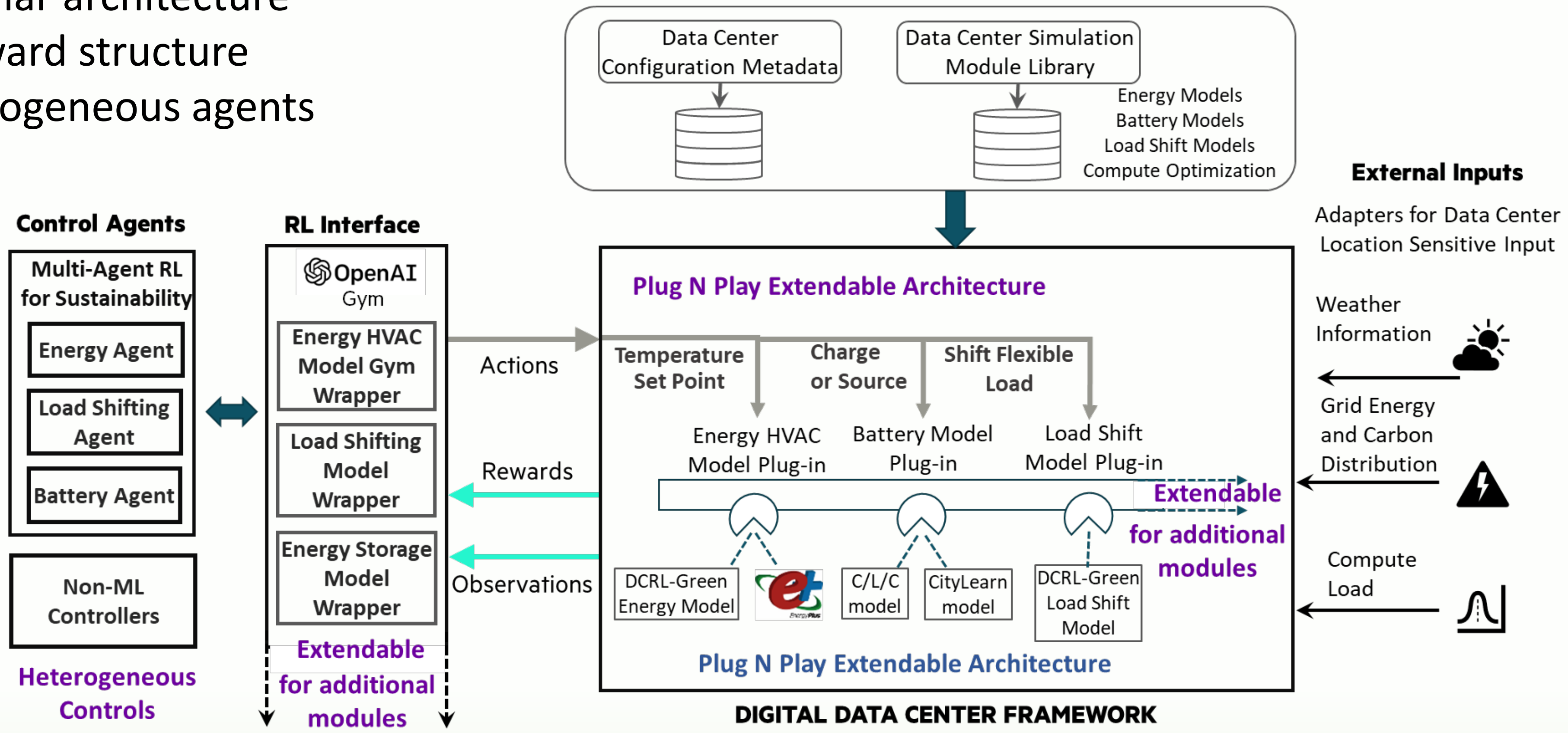
$$
(r_{LS}, r_{DC}, r_{BAT}) = \left(-(CFP_t + LS_{Penalty}), -(E_{hvac,t} + E_{it,t}), -(CFP_t)\right)
$$

$$
\begin{aligned}
R_{LS} &= \alpha * r_{LS} + (1-\alpha)/2 * r_{DC} + (1-\alpha)/2 * r_{BAT} \\
R_{DC} &= (1-\alpha)/2 * r_{LS} + \alpha * r_{DC} + (1-\alpha)/2 * r_{BAT} \\
R_{BAT} &= (1-\alpha)/2 * r_{LS} + (1-\alpha)/2 * r_{DC} + \alpha * r_{BAT}
\end{aligned}
$$

SustainDC Extendable Plug-in Framework

Framework for RL and other ML control for Data Centers

- Plug-and-play design
- Extendable modular architecture
- Customizable reward structure
- Support for heterogeneous agents

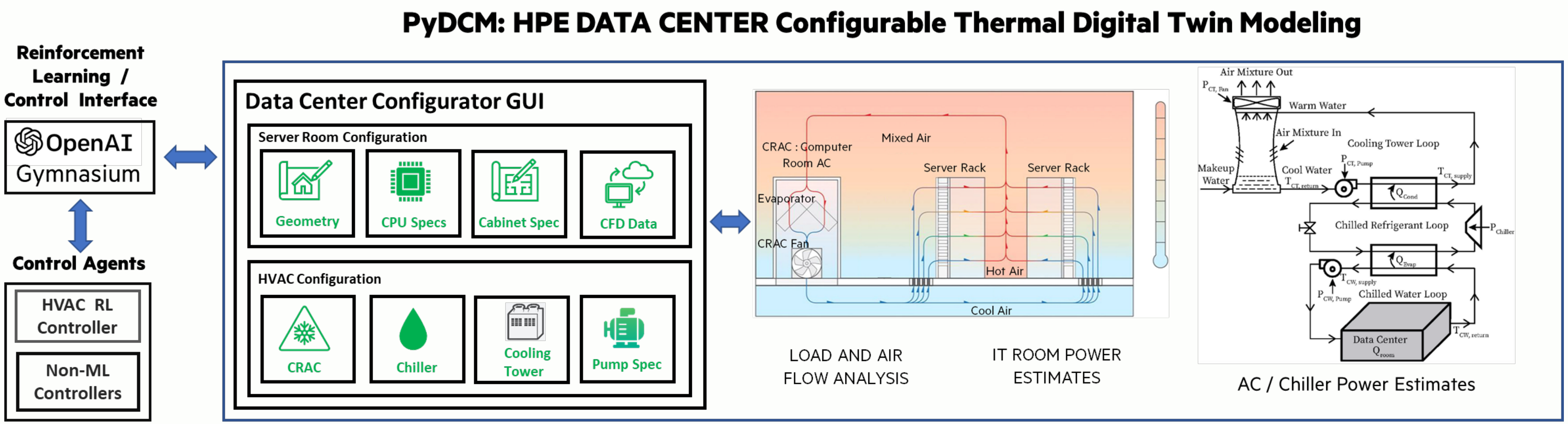


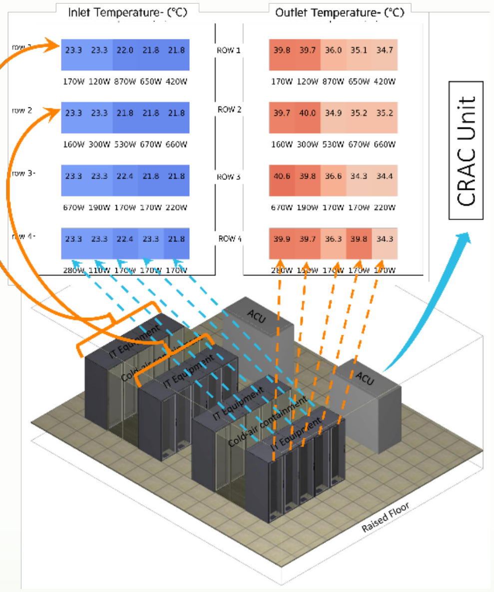
Custom Data Center Thermal Model

Current version support HVAC cooling

Will be extended to Liquid Cooling

- ❑ Modular, scalable, and highly customizable Data Center model suitable for testing a broad range of DC design parameters.
- ❑ Offers a unique Python-based platform that can be easily adapted for fast prototyping and control optimization by the energy / ML community.
- ❑ Serves as a powerful tool for the ML community to test ML / Reinforcement Learning optimization for DC models aimed at sustainability.

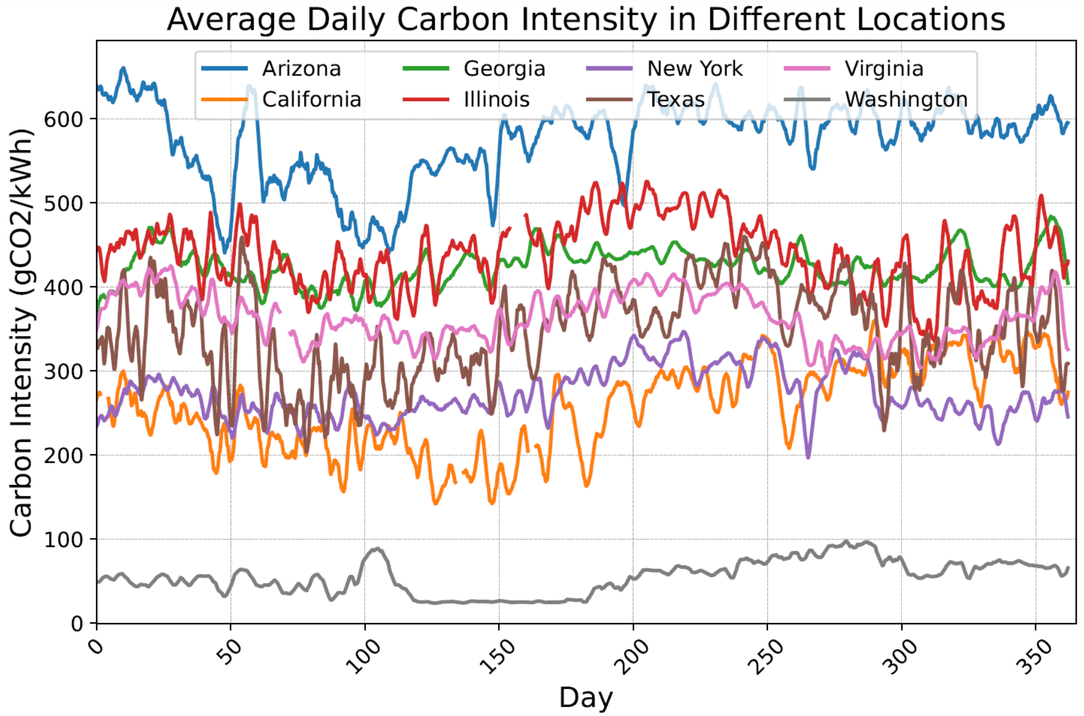




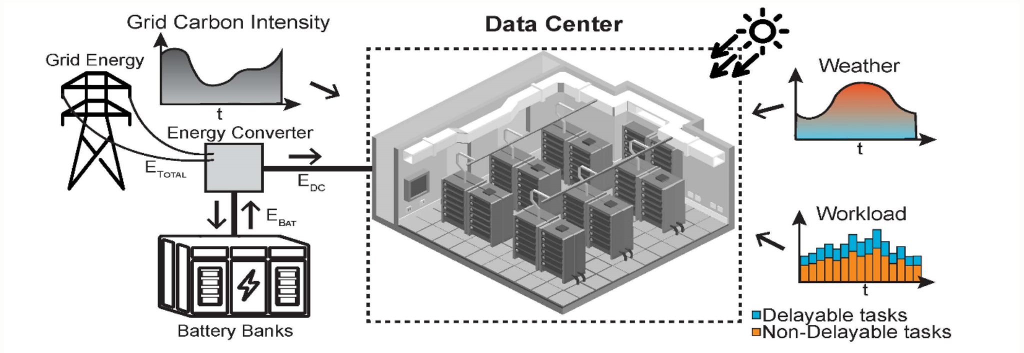
Experiments: Data Center Locations with variations

Location	Typical Weather	Carbon Emissions
Arizona	Hot, dry summers; mild winters	High avg CI, High variation
California	Mild, Mediterranean climate	Medium avg CI, Medium variation
Georgia	Hot, humid summers; mild winters	High avg CI, Medium variation
Illinois	Cold winters; hot, humid summers	High avg CI, Medium variation
New York	Cold winters; hot, humid summers	Medium avg CI, Medium variation
Texas	Hot summers; mild winters	Medium avg CI, High variation
Virginia	Mild climate, seasonal variations	Medium avg CI, Medium variation
Washington	Mild, temperate climate; wet winters	Low avg CI, Low variation

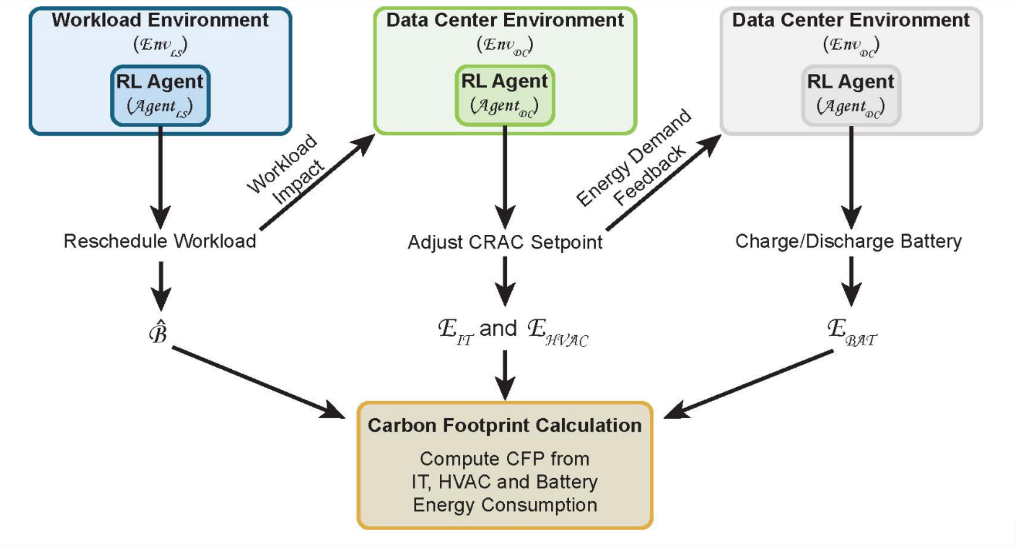
Table 7: Summary of Selected Locations with Typical Weather and Carbon Emissions Characteristics



Results: Support for various Multi-Agent RL algorithms

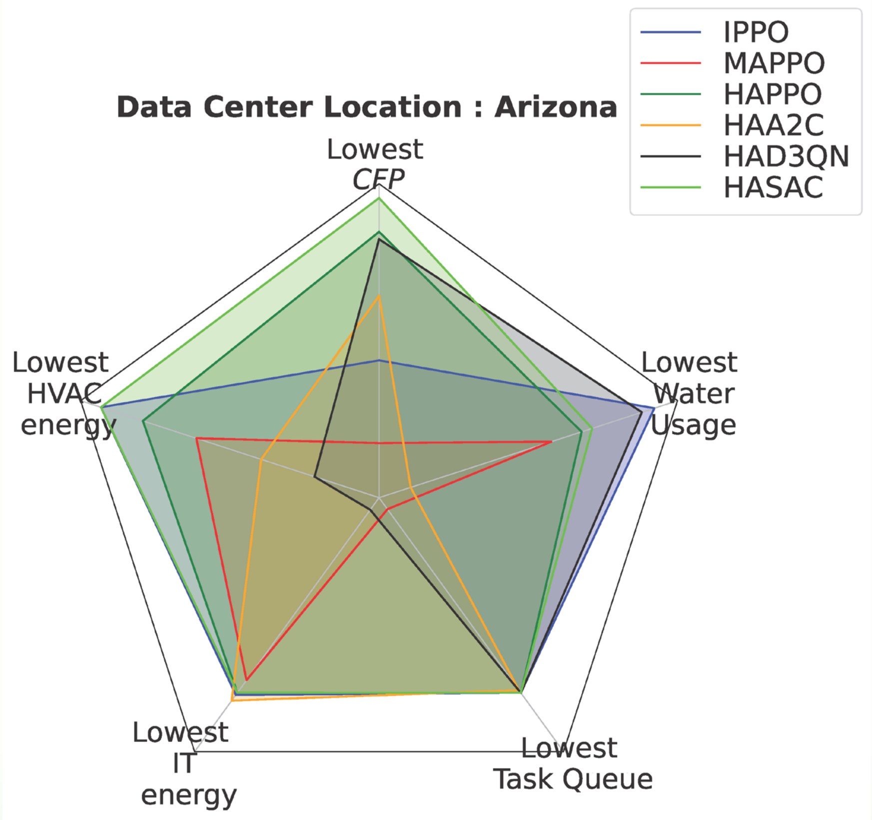


Physical System



Digital System

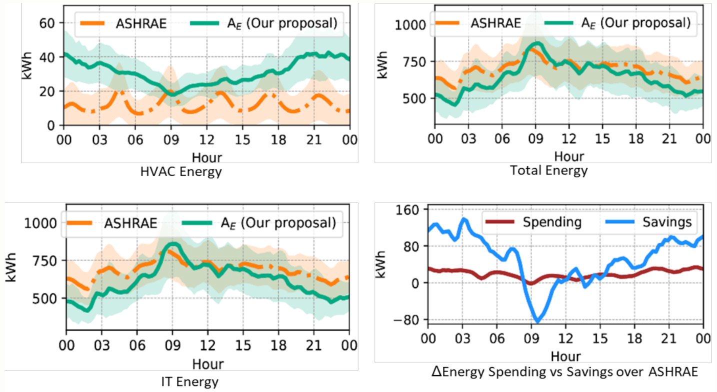
- Improvements in energy efficiency
- Carbon footprint reduction
- Lowering of water usage
- Latency of execution



Intuition on Optimization by various controls

NEURAL INFORMATION PROCESSING SYSTEMS

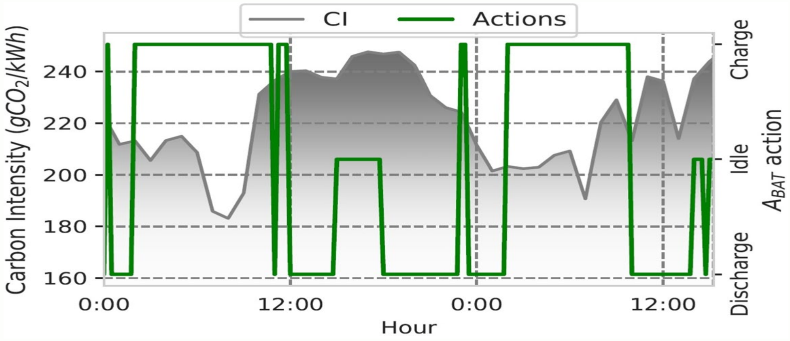
HVAC + IT Energy Optimization



HVAC Cooling Optimization over ASHRAE Controller

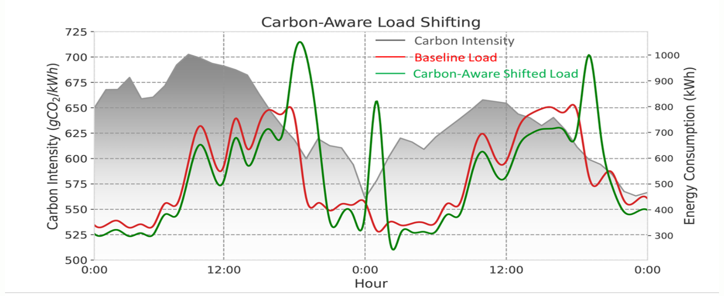
Investing more in **cooling energy** can significantly **reduce** overall **IT energy consumption.**

Energy Storage Optimization



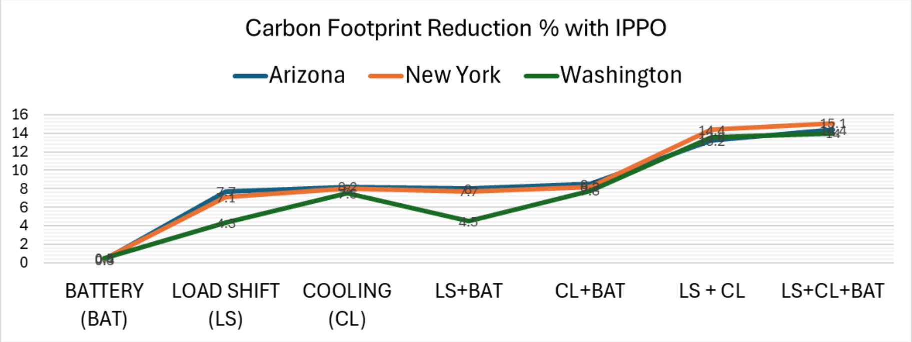
The **Battery** agent **stores energy** when the grid's **CI** is **low** and uses it when its **high**.

Load Shift Optimization



Carbon Intensity Aware Flexible Load Shifting

Flexible load is **shifted** to times when the grid's CI is **lower**.
It's not perfect because **predicting** the **CI** can be **challenging**.

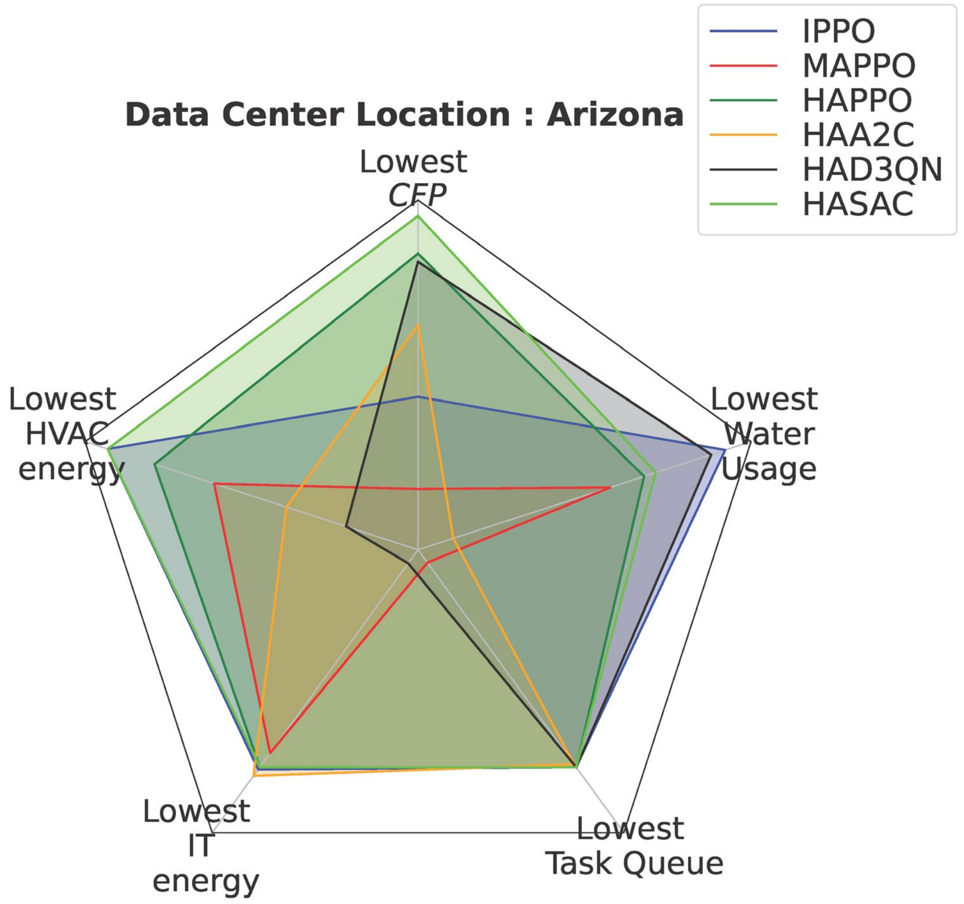


Cumulative Effect of Multiple Controls

Results: Support for various RL algorithms (Arizona)

Table 6: Multiagent RL framework evaluated on SustainDC for a data center located in Arizona (Average result over 5 runs)

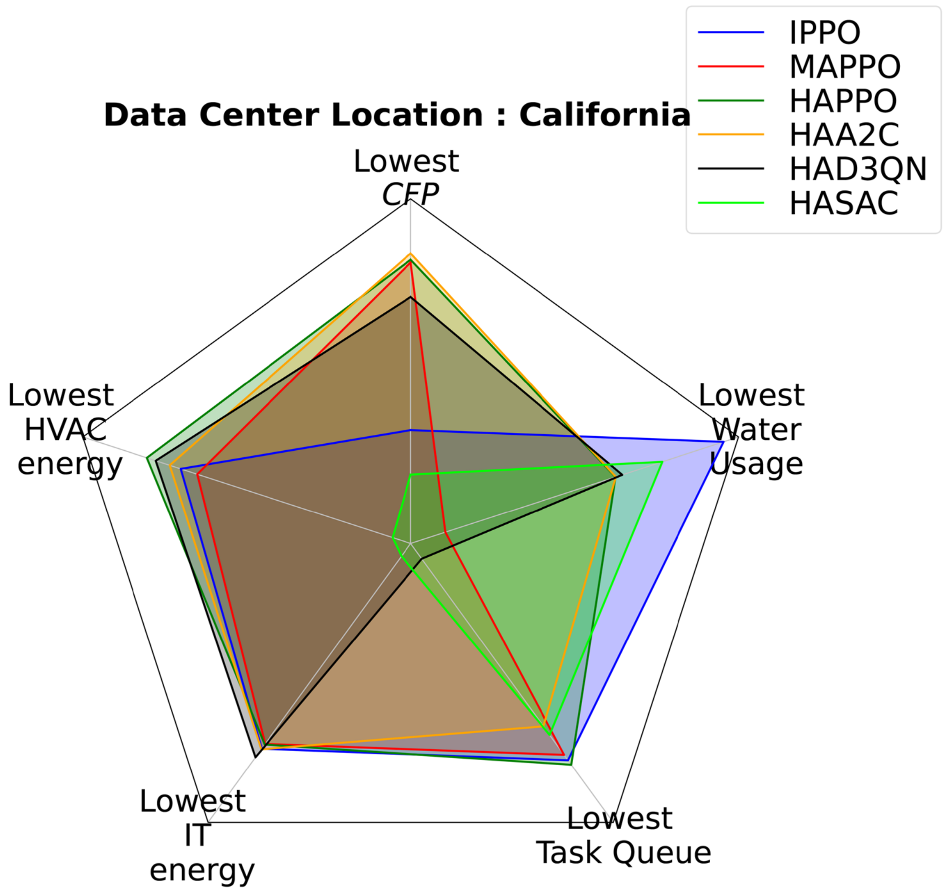
Evaluation Metric → / Algorithm ↓	CFP (kgCO2)	HVAC Energy (kwh)	IT Energy (kwh)	Task Queue	Water Usage (litre)
IPPO	408.7	380.8	934.8	0.60	30251.6
MAPPO	410.8	383.3	947.5	502.4	31289.6
HAPPO	405.5	381.9	936.6	0.26	30983.7
HAA2C	407.1	385.0	929.9	7.54	32706.3
HAD3QN	405.6	386.4	1094.0	0.0051	30377.3
HASAC	404.6	380.8	936.7	0.54	30878.7



Results: Support for various RL algorithms (California)

Table 5: Multiagent RL framework evaluated on SustainDC for a data center located in California (Average result over 5 runs)

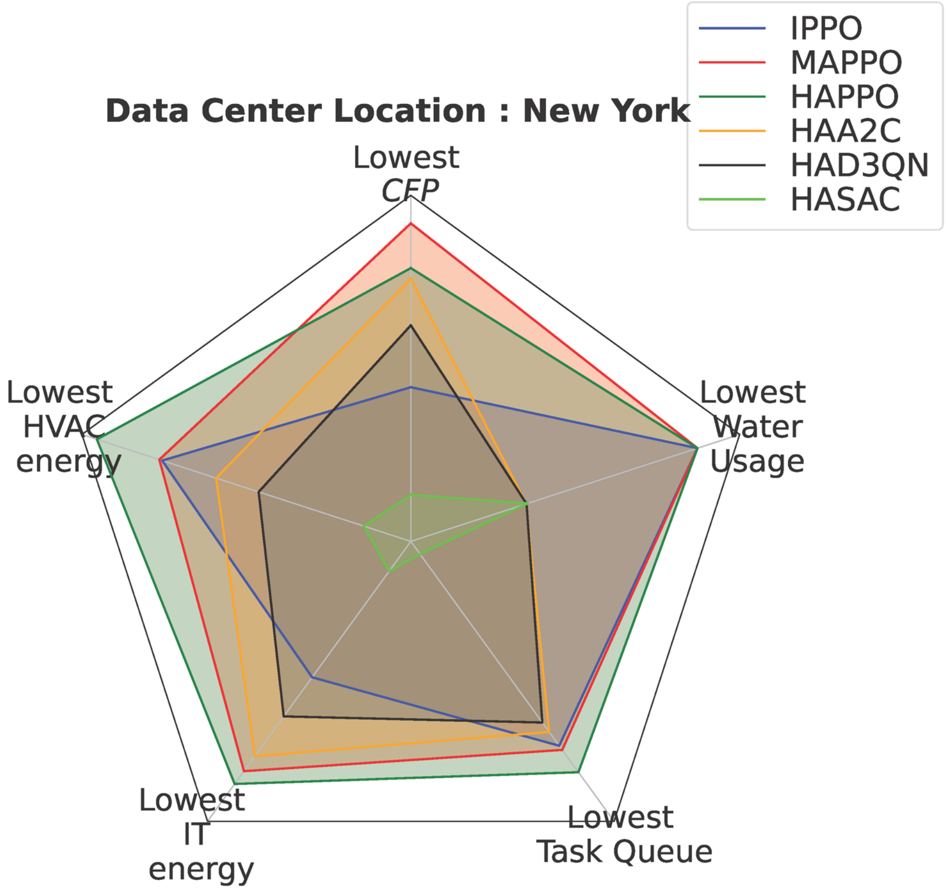
Evaluation Metric → Algorithm ↓	CFP (kgCO2)	HVAC Energy (kwh)	IT Energy (kwh)	Task Queue	Water Usage (litre)
IPPO	170.0	384.3	933.8	12.9	28141.4
MAPPO	159.3	388.2	936.1	19.5	33289.3
HAPPO	159.1	376.3	935.8	74.9	30141.8
HAA2C	158.7	381.7	933.5	54.1	30135.4
HAD3QN	161.5	378.4	929.6	25.8	30017.4
HASAC	172.9	434.4	1027.0	43.8	29277.5



Results: Support for various RL algorithms (New York)

Table 3: Multiagent RL framework evaluated on SustainDC for a data center located in New York (Average result over 5 runs)

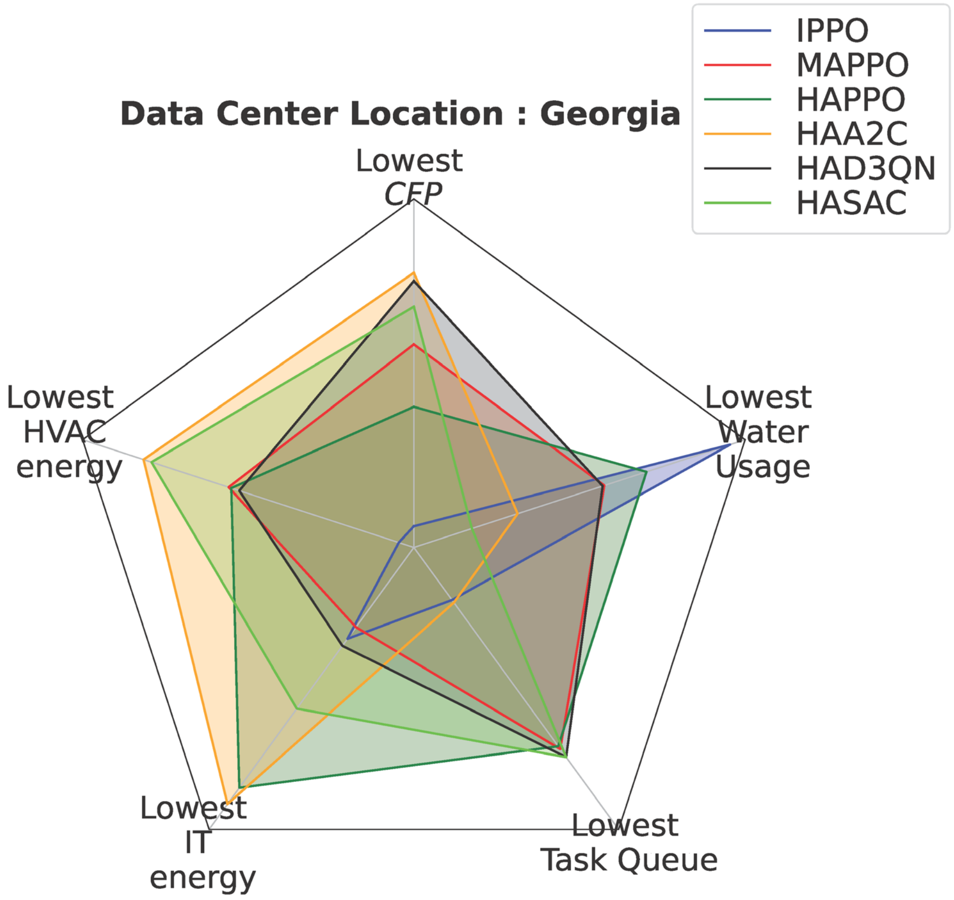
Evaluation Metric → Algorithm ↓	CFP (kgCO2)	HVAC Energy (kwh)	IT Energy (kwh)	Task Queue	Water Usage (litre)
IPPO	179.6	417.1	945.9	20.9	446.2
MAPPO	176.4	417.0	932.7	19.6	446.2
HAPPO	177.3	414.8	930.9	12.8	441.9
HAA2C	177.5	419.0	934.8	25.2	14977.1
HAD3QN	178.4	420.5	940.4	28.0	14950.9
HASAC	181.7	424.2	960.8	79.7	14842.4



Results: Support for various RL algorithms (Georgia)

Table 4: Multiagent RL framework evaluated on SustainDC for a data center located in Georgia (Average result over 5 runs)

Evaluation Metric → Algorithm ↓	CFP (kgCO2)	HVAC Energy (kwh)	IT Energy (kwh)	Task Queue	Water Usage (litre)
IPPO	265.4	376.7	935.4	6.8	31773.5
MAPPO	263.4	370.3	935.9	0.35	31949.9
HAPPO	264.1	370.4	929.0	0.47	31890.7
HAA2C	262.7	367.1	928.3	6.6	32071.5
HAD3QN	262.8	370.7	935.1	0.0	31952.2
HASAC	263.0	367.4	932.4	0.0	32135.7

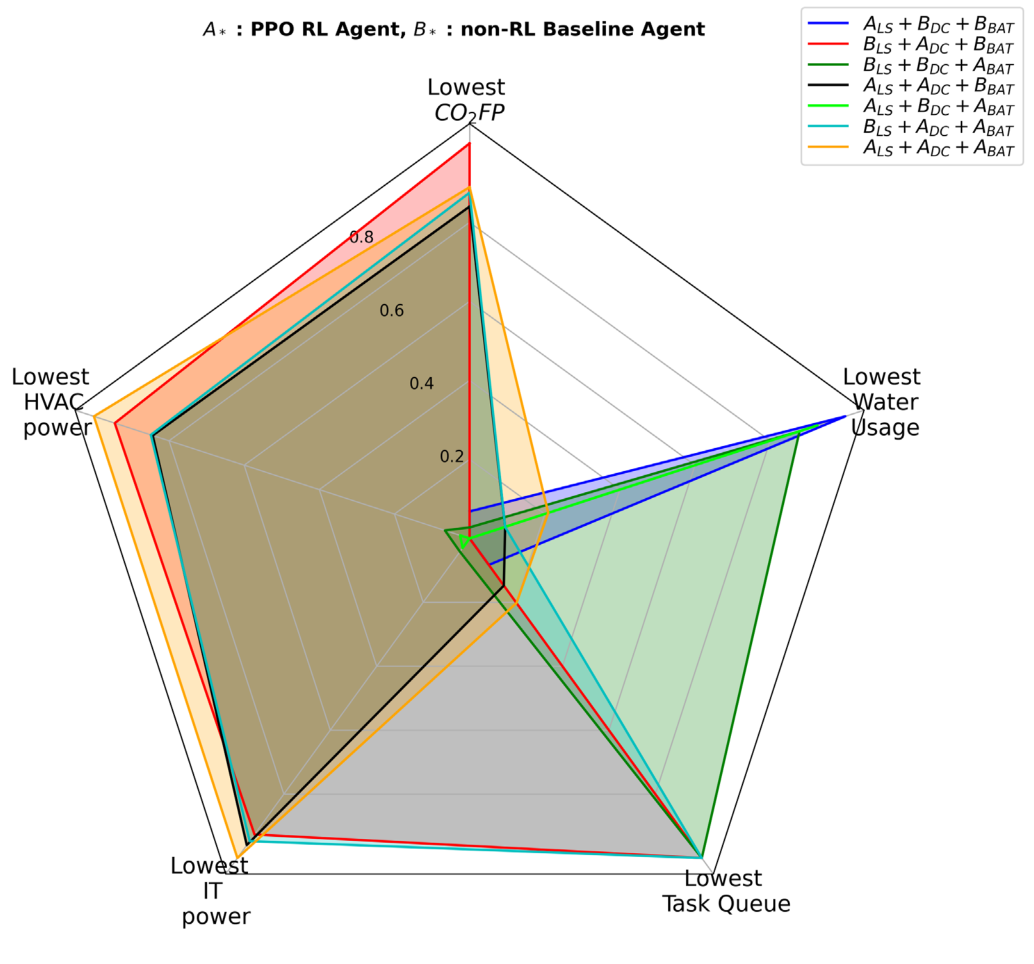


Results: Benefits of Multi-Agent Control

- Mixture of RL and Baseline agents
- Improvements in energy efficiency
- Carbon footprint reduction
- Lowering of water usage
- Latency of execution

Table 1: Performance with respect to evaluation metrics on single and multiple RL agent baselines. A_* : $RL\ agent\ B_*$: $non-RL\ baseline\ agent$

Evaluation Metric → Algorithm ↓	CFP (kgCO2)	HVAC Energy (kwh)	IT Energy (kwh)	Task Queue	Water Usage (litre)
1:$A_{LS}+B_{DC}+B_{BAT}$	167.61	391.6	1033.8	0.52	10433.46
2:$B_{LS}+A_{DC}+B_{BAT}$	153.56	372.9	944.5	0.0	10930.77
3:$B_{LS}+B_{DC}+A_{BAT}$	168.22	390.3	1029.8	0.0	10493.95
4: $A_{LS}+A_{DC}+B_{BAT}$	155.97	374.9	941.3	0.48	10883.73
5:$A_{LS}+B_{DC}+A_{BAT}$	168.64	391.1	1030.9	0.56	10470.43
6:$B_{LS}+A_{DC}+A_{BAT}$	155.44	374.8	942.5	0	10883.73
7:$A_{LS}+A_{DC}+A_{BAT}$	155.23	371.8	937.4	0.45	10826.61



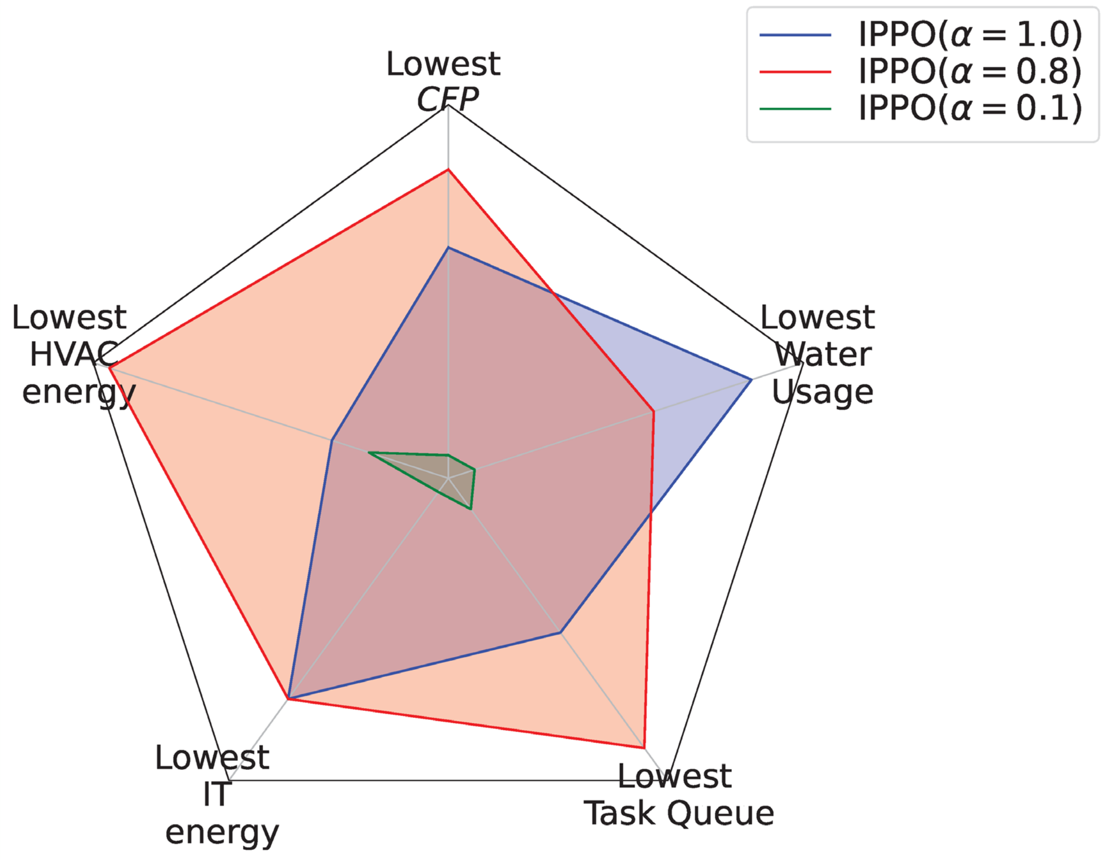
Results: Collaborative multi-agent hyperparameter tuning

NEURAL INFORMATION PROCESSING SYSTEMS

Table 2: IPPO evaluated on SustainDC with different values of collaborative reward coefficient α (Average result over 12 runs)

Evaluation Metric → / Algorithm ↓	CFP (kgCO2)	HVAC Energy (kwh)	IT Energy (kwh)	Task Queue	Water Usage (litre)
IPPO($\alpha = 1.0$)	176.3	415.2	932.8	12.5	445.6
IPPO($\alpha = 0.8$)	176.2	414.6	932.8	9.5	445.8
IPPO($\alpha = 0.1$)	176.4	415.3	932.9	15.7	446.2

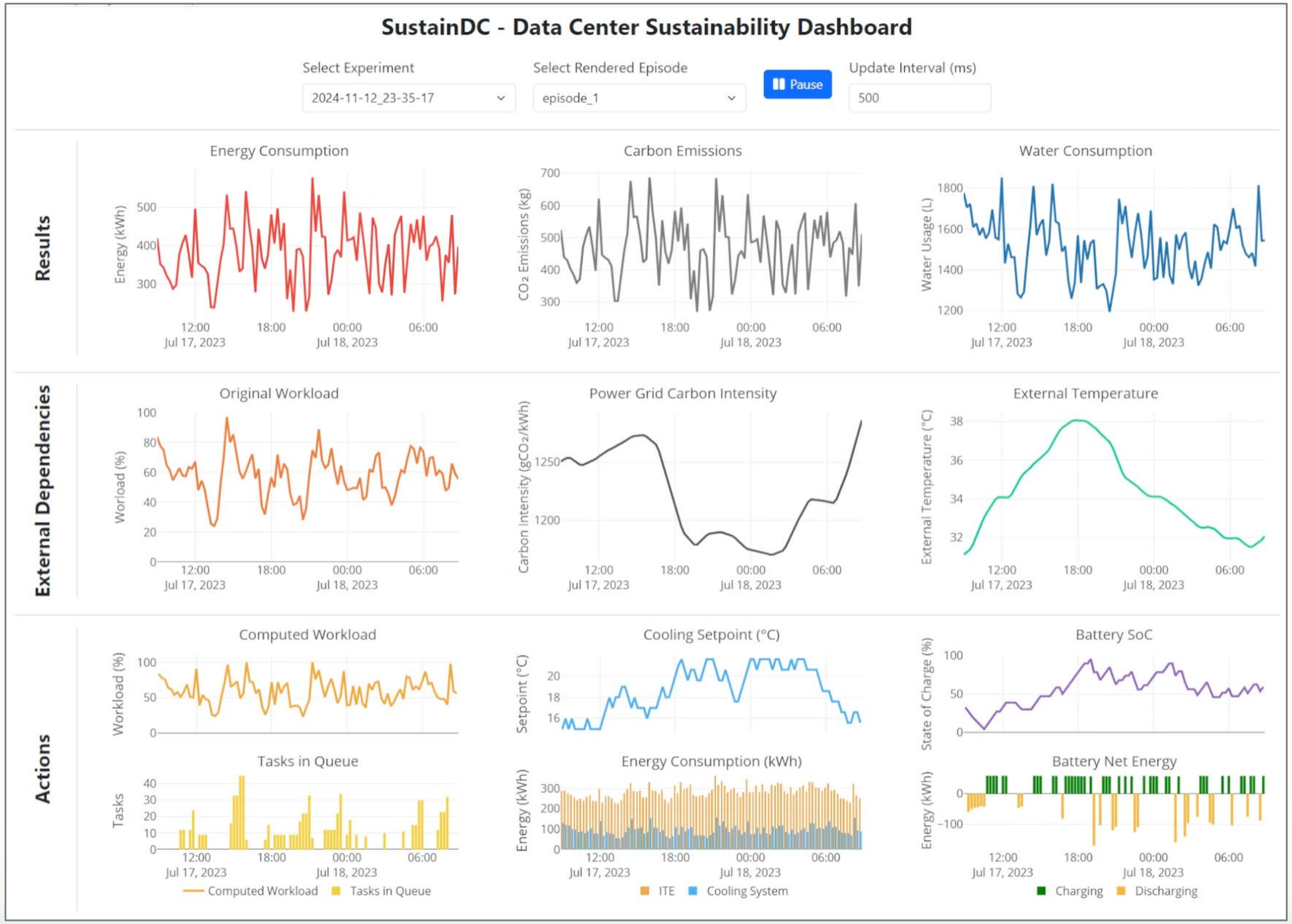
$$R_{LS} = \alpha * r_{LS} + (1-\alpha)/2 * r_{DC} + (1-\alpha)/2 * r_{BAT}$$
$$R_{DC} = (1-\alpha)/2 * r_{LS} + \alpha * r_{DC} + (1-\alpha)/2 * r_{BAT}$$
$$R_{BAT} = (1-\alpha)/2 * r_{LS} + (1-\alpha)/2 * r_{DC} + \alpha * r_{BAT}$$



Heterogeneous Multi-objective RL Sustainability Benchmark

NEURAL INFORMATION PROCESSING SYSTEMS

This is the first comprehensive Multi-agent Multi-objective RL Benchmark for evaluating RL Algorithms for Sustainability with multiple Internal and External Dependencies

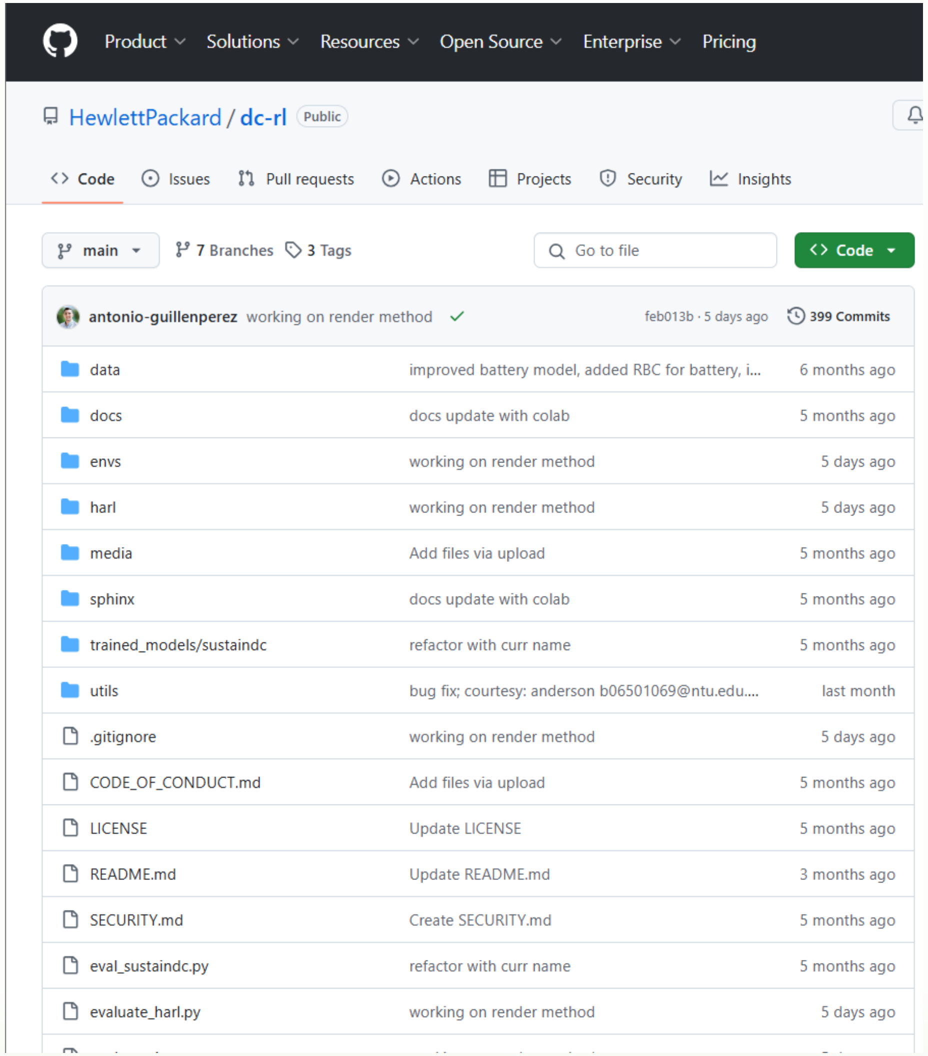


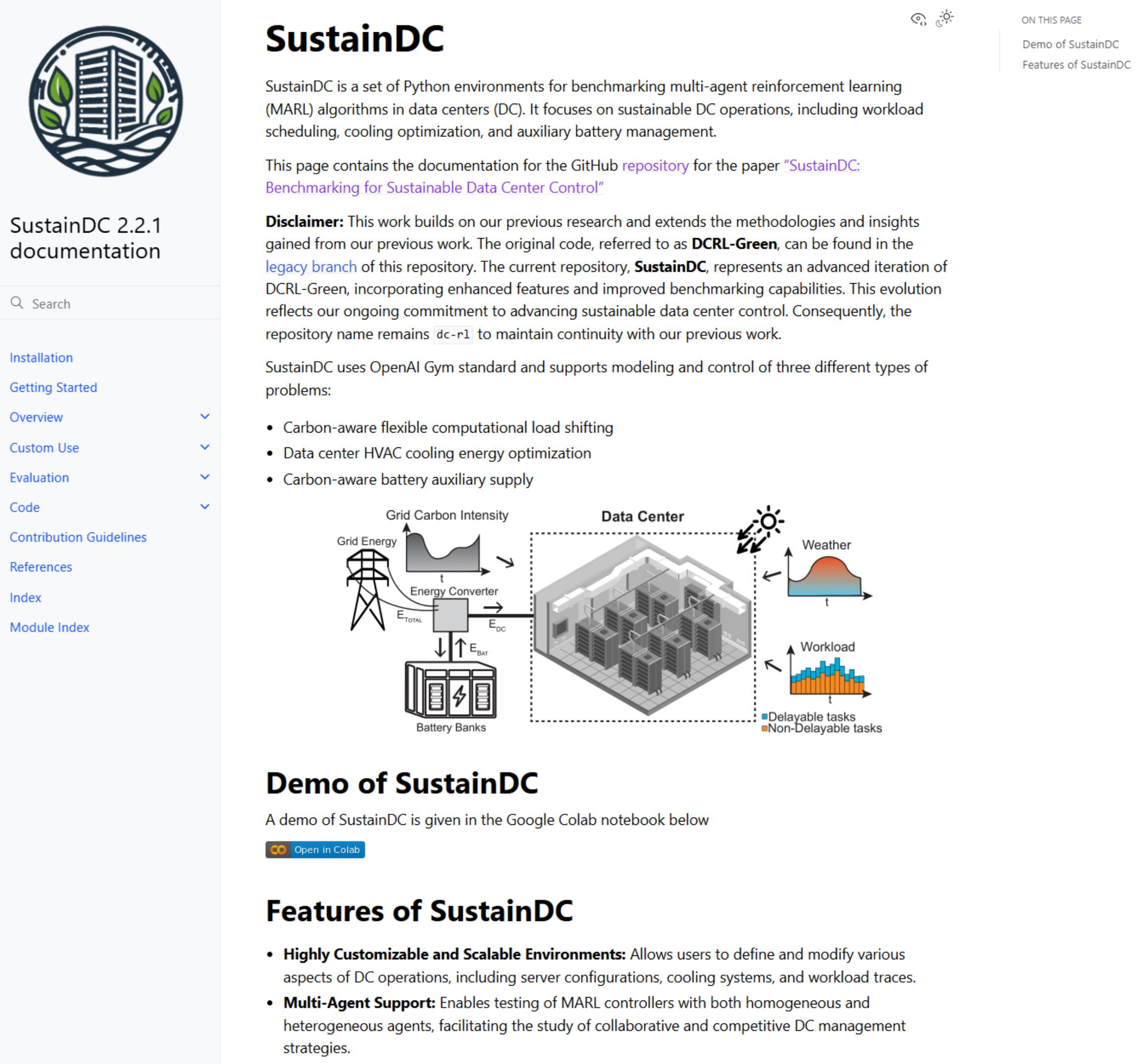
GitHub

https://github.com/HewlettPackard/dc-rl

Multi-agent Multi-objective RL Sustainability Benchmark

NEURAL INFORMATION PROCESSING SYSTEMS



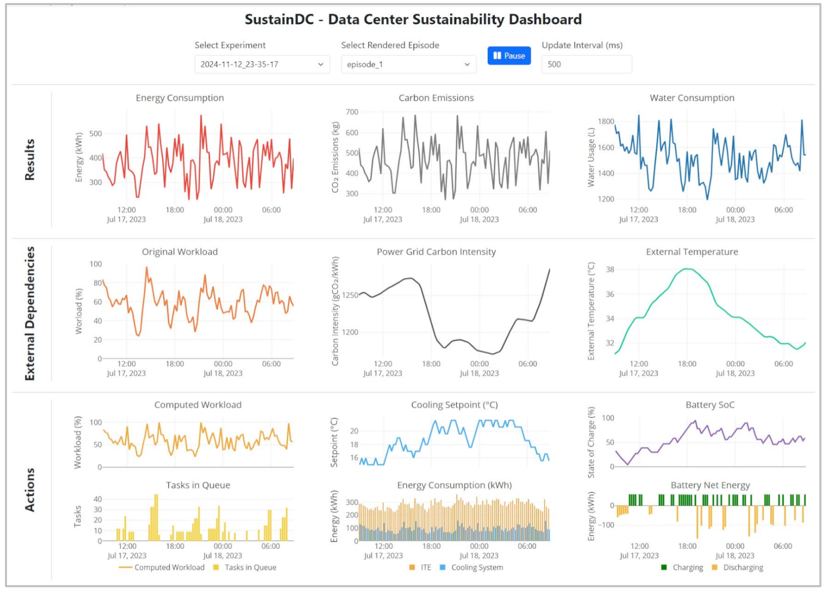


Hewlett Packard Enterprise

NeurIPS 2024

NEURAL INFORMATION PROCESSING SYSTEMS

Thank You



GitHub

OpenReview

Avisek Naug

Antonio Guillen

Ricardo Luna

Vineet Gundecha

Cullen Bash

Sahand Ghorbanpour

Sajad Mousavi

Ashwin R Babu

Soumyendu Sarkar

ML Research
@Hewlett Packard Labs

soumyendu.sarkar@hpe.com