WaterWise: Co-optimizing Carbon- and Water-Footprint Toward Environmentally Sustainable Cloud Computing




Yankai Jiang | Rohan Basu Roy | Raghavendra Kanakagiri | Devesh Tiwari

Northeastern University
The University of Utah
भारतीय प्रौद्योगिकी संस्थान तिरुपति
IIT Tirupati

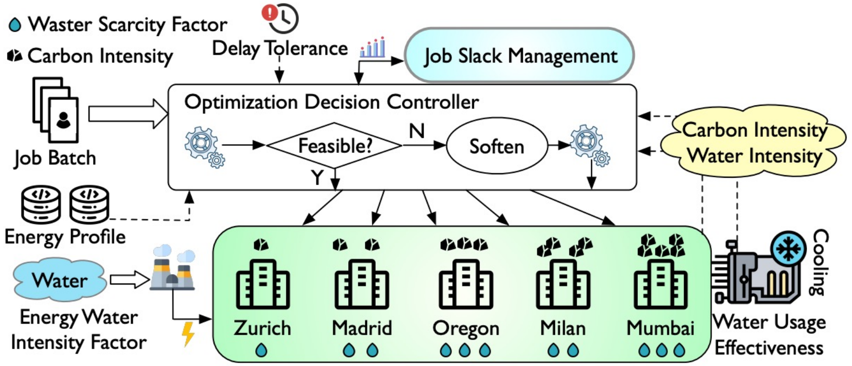
WaterWise: Executive Summary

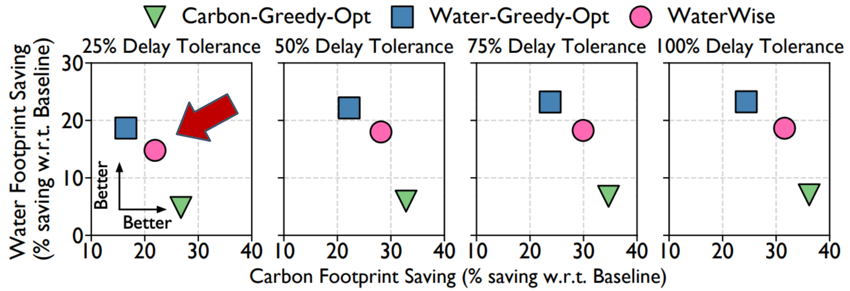
- ✓ **A Novel Sustainability-Focused Job Scheduler** Carbon and water footprint savings are often at odds with each other, and require intelligent exploitation of temporal and spatial opportunities.

- ✓ **Key Ideas of WaterWise** Delay tolerance to maximize the carbon and water footprint savings for parallel workloads. Comprehensive modeling of water footprint of geographically distributed data centers.

- ✓ **Built-in Awareness of Water Scarcity Factor** A water drop may not have equal value at different geographical locations.

Mixed Integer Linear Programming Formulation



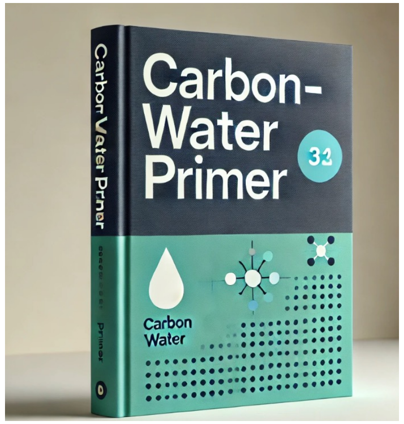


Before sharing the lessons learned from designing WaterWise…

Before sharing the lessons learned from designing WaterWise…

… a primer on carbon and water footprint models

WaterWise's comprehensive water footprint modeling



Carbon Footprint Modeling

$$CO_{2,j} = CO_{2,j}^{operational} + CO_{2,j}^{embodied}$$

$$CO_{2,j} = E_j \cdot CO_2^{Intensity} + \frac{t_j}{T_{lifetime}} \cdot CO_{2,server}^{embodied}$$

Embodied and Operational Carbon

Accounting for carbon emissions of energy sources during operations (carbon intensity) and carbon emission during manufacturing.

Carbon Footprint Modeling

$$CO_{2,j} = CO_{2,j}^{operational} + CO_{2,j}^{embodied}$$

$$CO_{2,j} = E_j \cdot CO_2^{Intensity} + \frac{t_j}{T_{lifetime}} \cdot CO_{2,server}^{embodied}$$

Embodied and Operational Carbon

Accounting for carbon emissions of energy sources during operations (carbon intensity) and carbon emission during manufacturing.

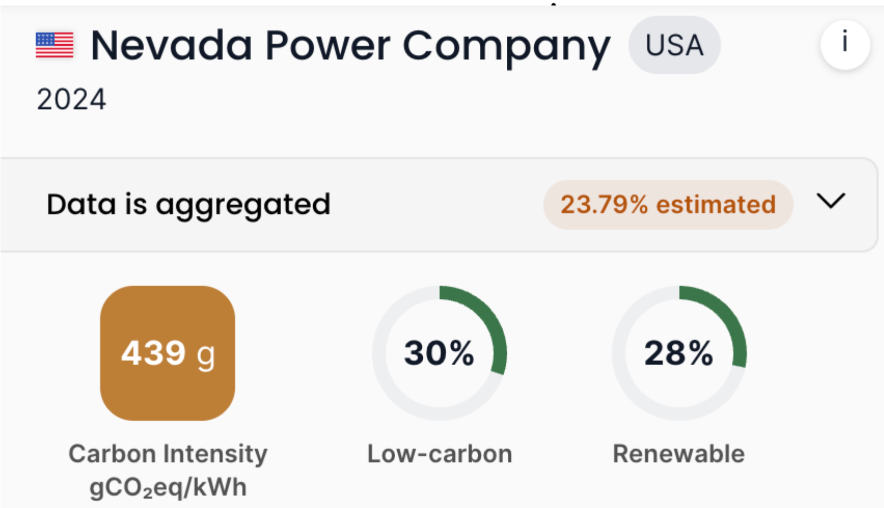
Carbon Footprint Modeling

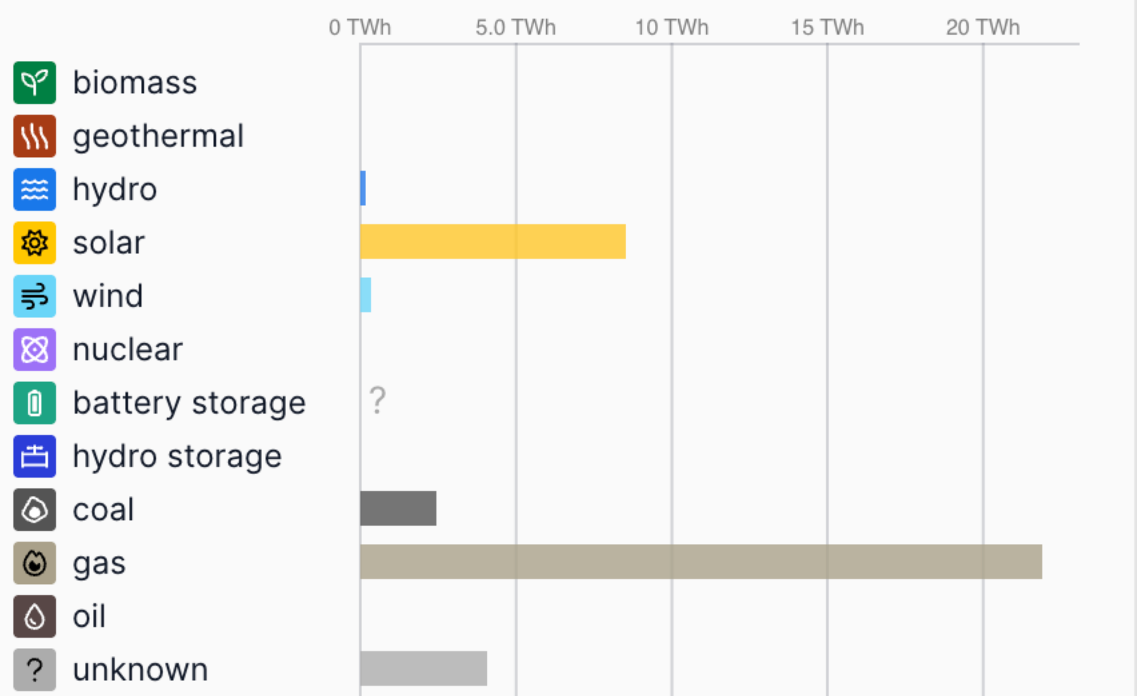
$$CO_{2,j} = CO_{2,j}^{\text{operational}} + CO_{2,j}^{\text{embodied}}$$

$$CO_{2,j} = E_j \cdot CO_2^{\text{Intensity}} + \frac{t_j}{T_{\text{lifetime}}} \cdot CO_{2,\text{server}}^{\text{embodied}}$$

Embodied and Operational Carbon

Accounting for carbon emissions of energy sources during operations (carbon intensity) and carbon emission during manufacturing.





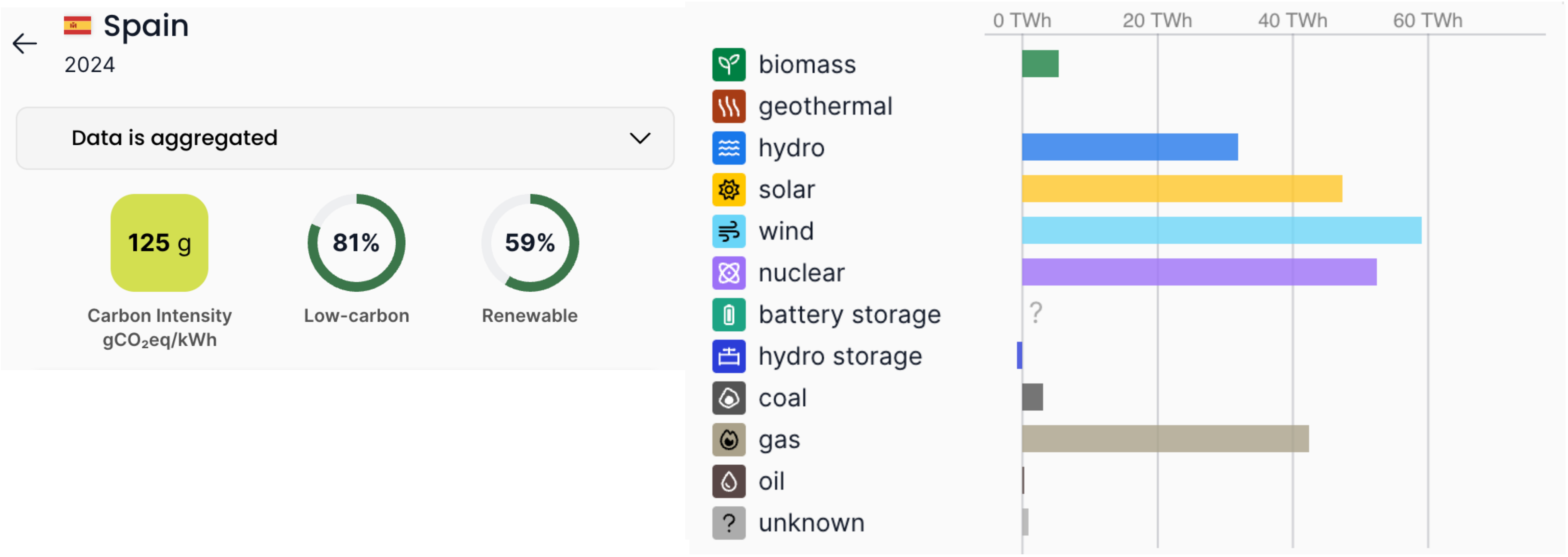
Carbon Footprint Modeling

$$CO_{2,j} = CO_{2,j}^{\text{operational}} + CO_{2,j}^{\text{embodied}}$$

$$CO_{2,j} = E_j \cdot CO_2^{\text{Intensity}} + \frac{t_j}{T_{\text{lifetime}}} \cdot CO_{2,\text{server}}^{\text{embodied}}$$

Embodied and Operational Carbon

Accounting for carbon emissions of energy sources during operations (carbon intensity) and carbon emission during manufacturing.



WaterWise Water Footprint Modeling of Data Centers

Water Scarcity Factor

Operational Water Footprint (Onsite)

Operational Water Footprint (Offsite)

Embodied Water Footprint

WaterWise Water Footprint Modeling of Data Centers

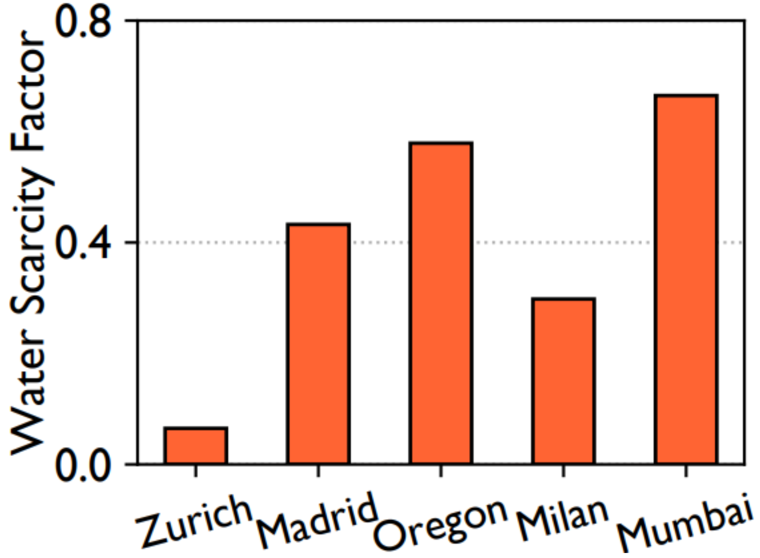
Water Scarcity Factor

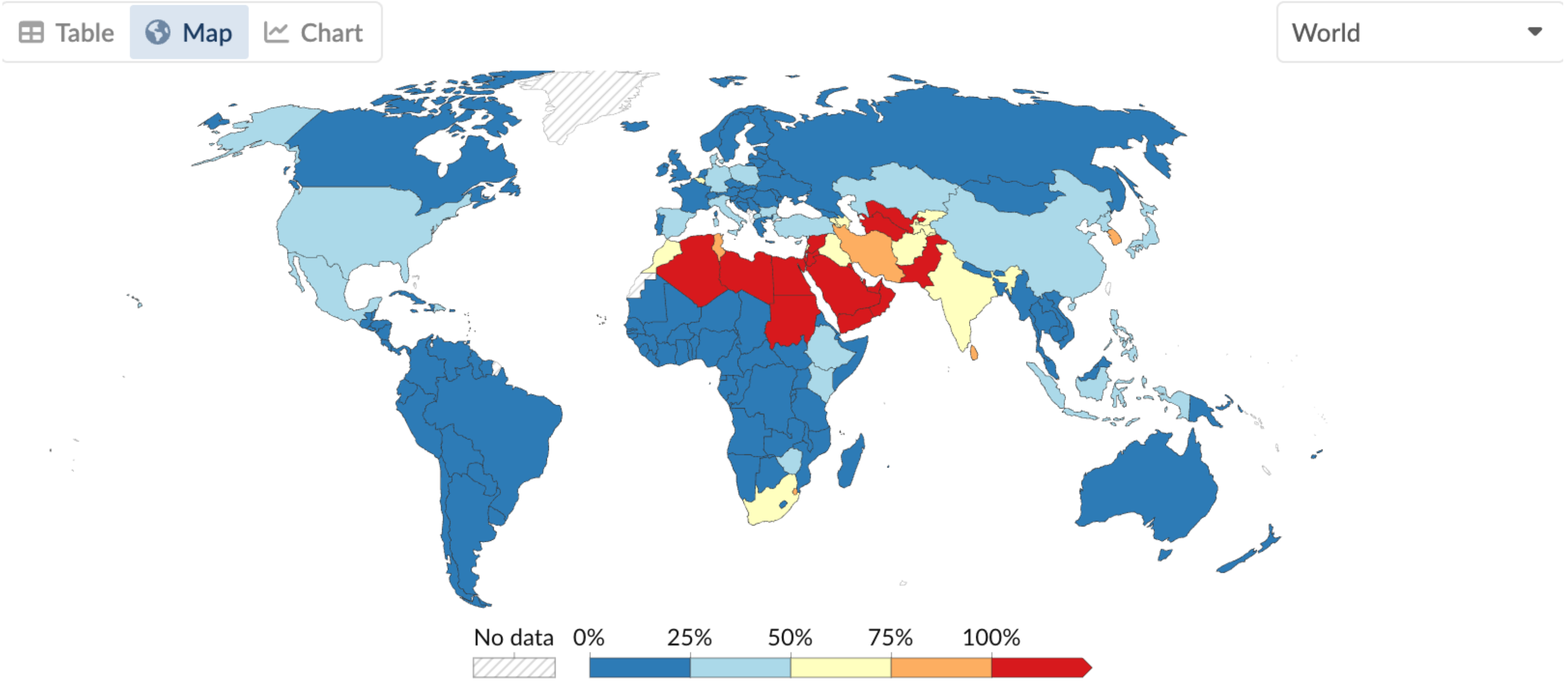
Operational Water Footprint (Onsite)

Operational Water Footprint (Offsite)

Embodied Water Footprint

Water Scarcity Factor (WSF)





The Water Scarcity Factor (WSF) serves to gauge the degree of water stress in specific regions, and measure the availability of freshwater relative to demand (a higher WSF indicates that the region is more scarce).

WaterWise Water Footprint Modeling of Data Centers

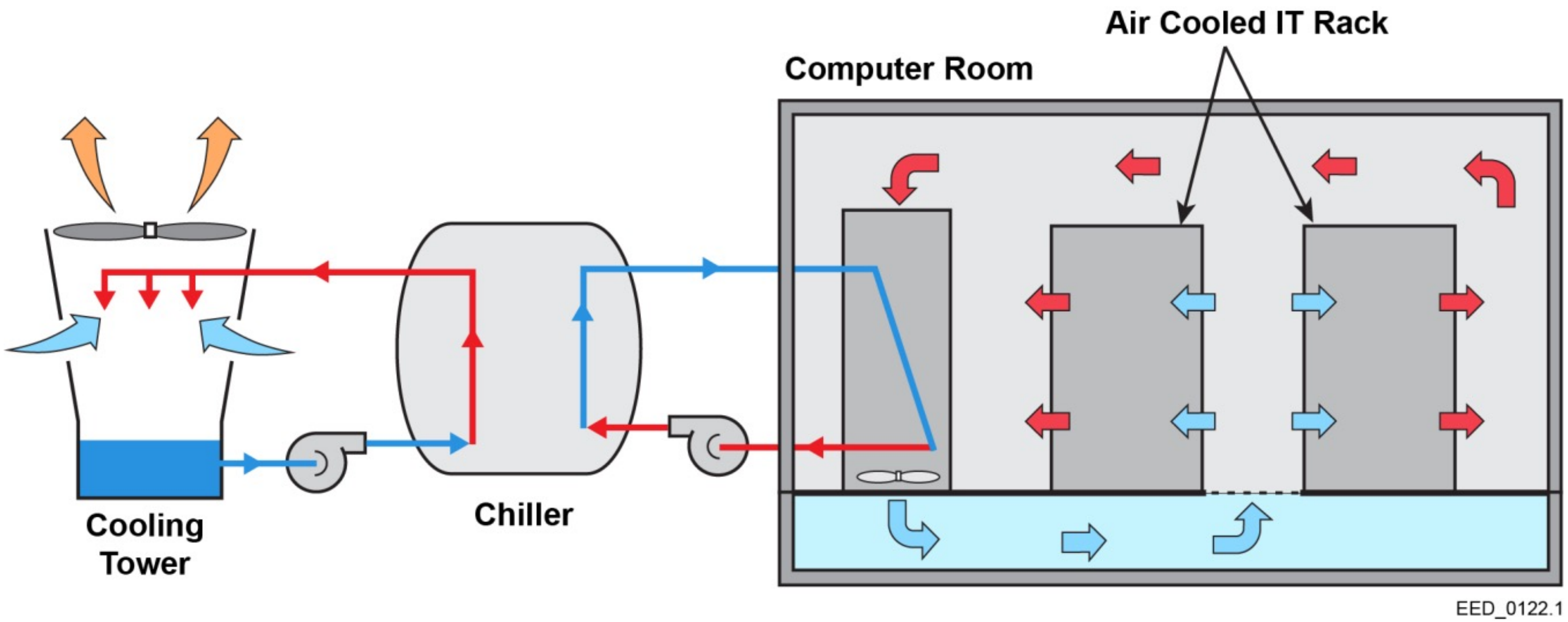
Water Scarcity Factor

Operational Water Footprint (Onsite)

Operational Water Footprint (Offsite)

Embodied Water Footprint

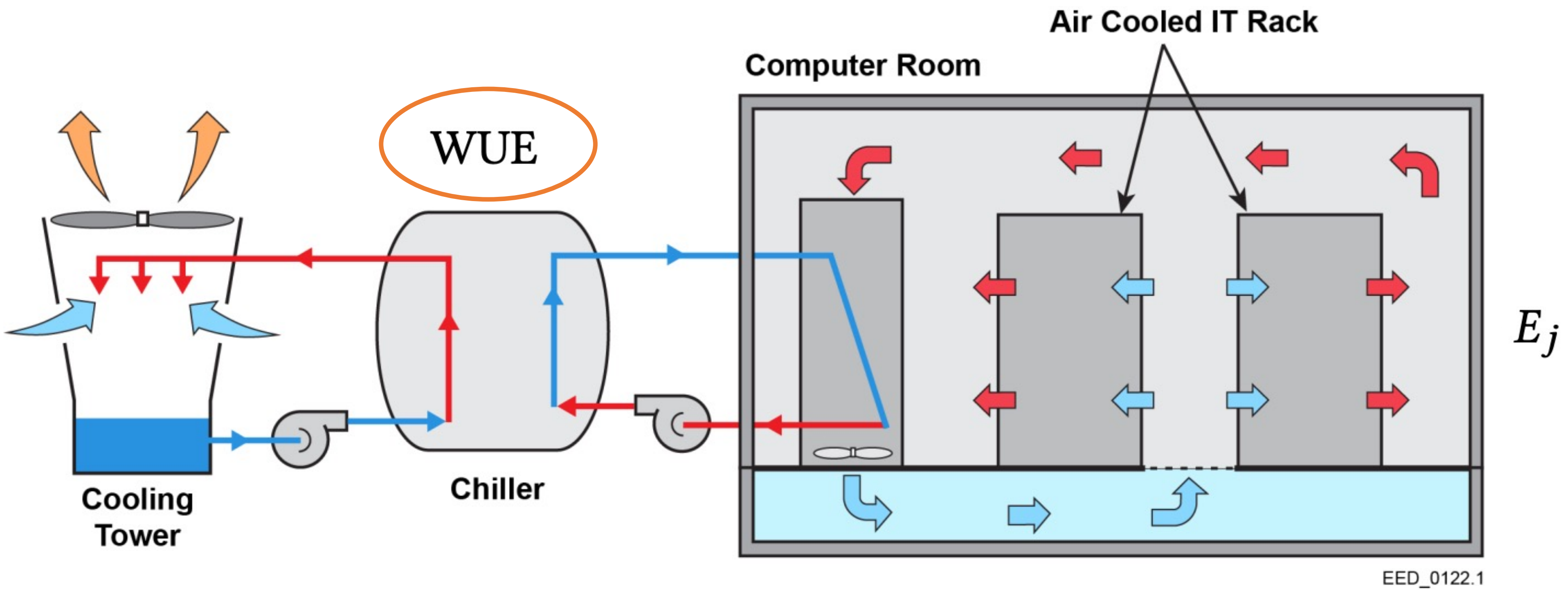
Operational Water Footprint (Onsite)



The onsite water footprint refers to the amount of water evaporated and dissipated during heat transfer and blowdown in the cooling process.

Reference: Cooling Water Efficiency Opportunities for Federal Data Centers, DOE

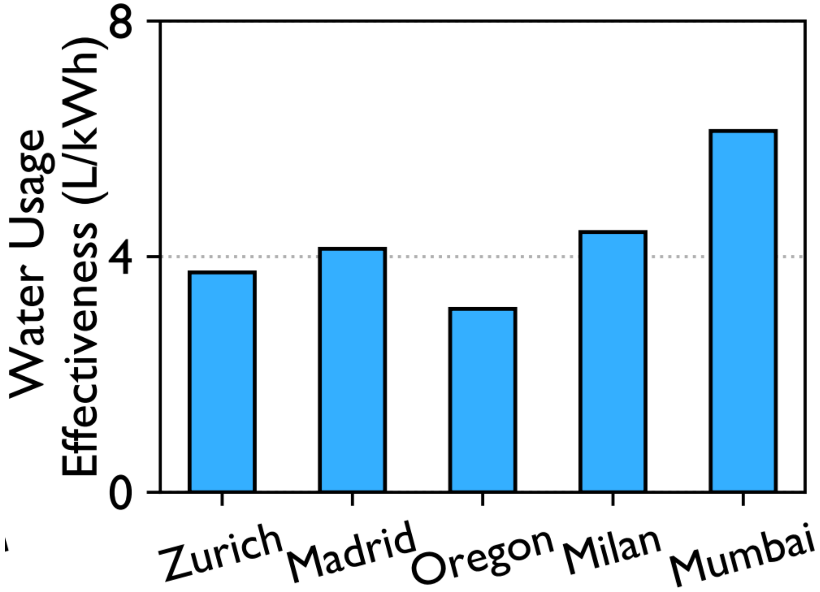
Operational Water Footprint (Onsite)

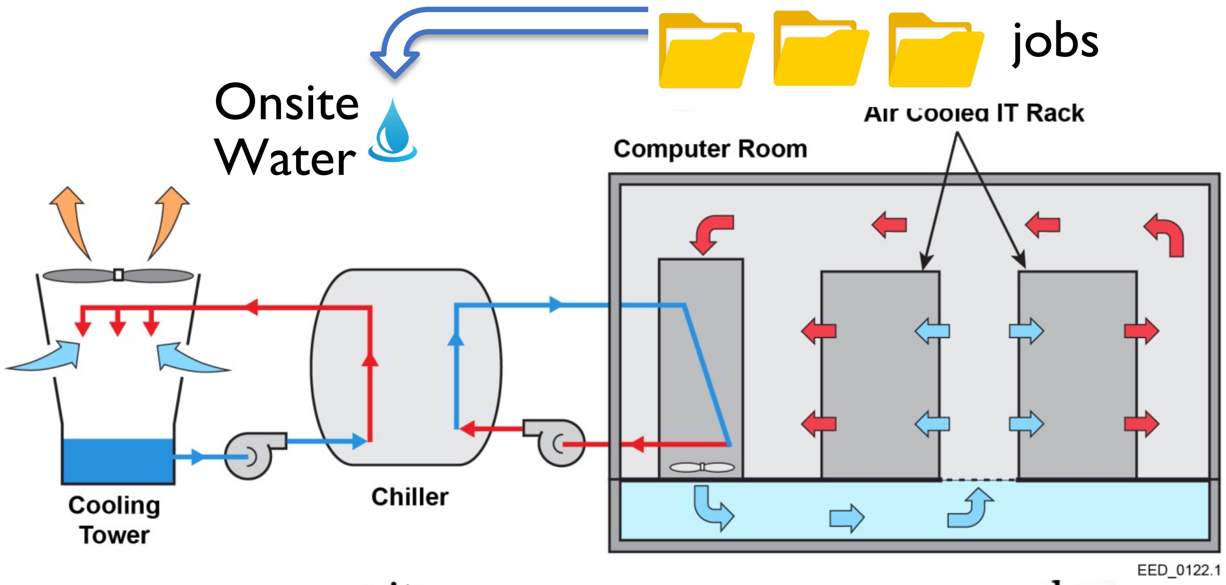


The onsite water footprint refers to the amount of water evaporated and dissipated during heat transfer and blowdown in the cooling process.

Reference: Cooling Water Efficiency Opportunities for Federal Data Centers, DOE

Operational Water Footprint (Onsite)





$$H_2O_j^{\mathrm{onsite}} = E_j \times \mathrm{WUE} \times (1 + \mathrm{WSF}_r^{\mathrm{dc}})$$

The Water Usage Effectiveness (WUE) metric quantifies the water required to dissipate heat per unit of energy generated, measured in L/kWh (lower is better).

Reference: Cooling Water Efficiency Opportunities for Federal Data Centers, DOE

WaterWise Water Footprint Modeling of Data Centers

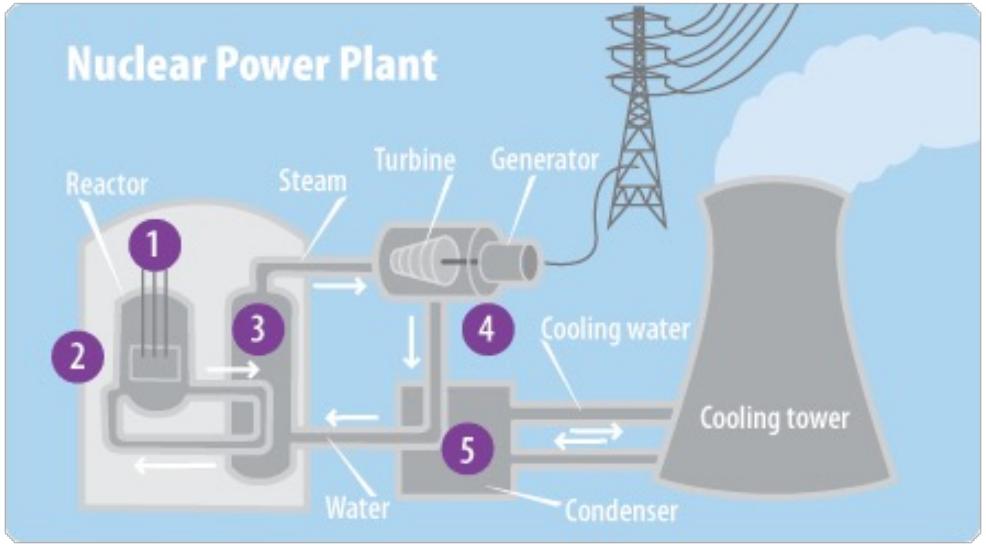
Water Scarcity Factor

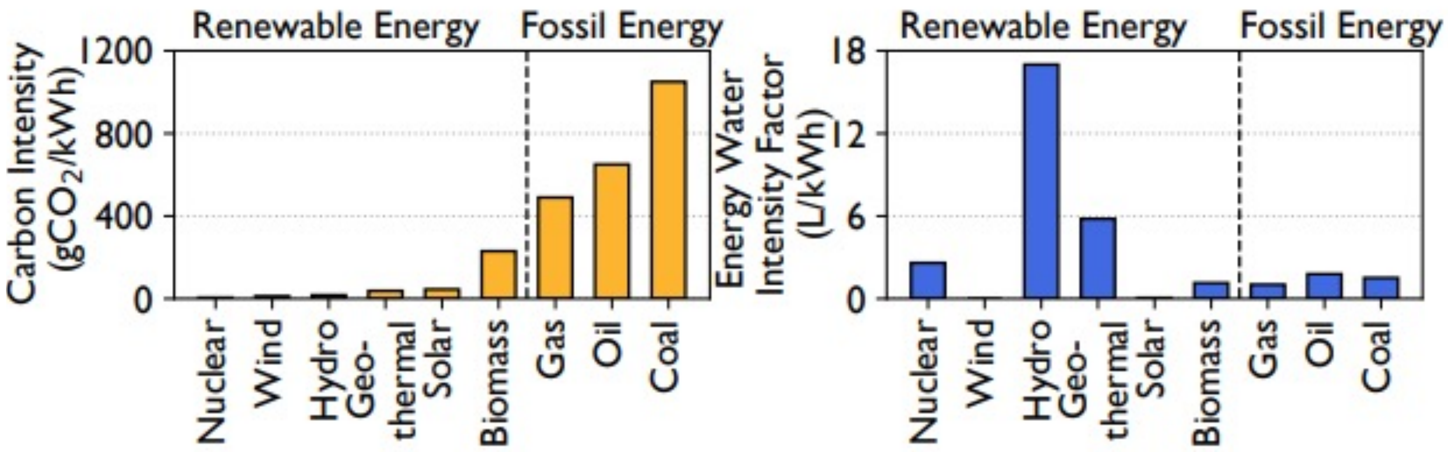
Operational Water Footprint (Onsite)

Operational Water Footprint (Offsite)

Embodied Water Footprint

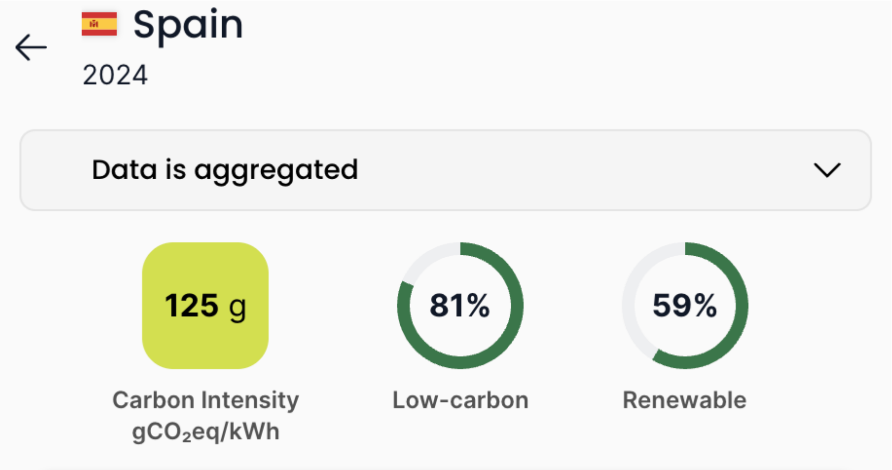
Operational Water Footprint (Offsite)

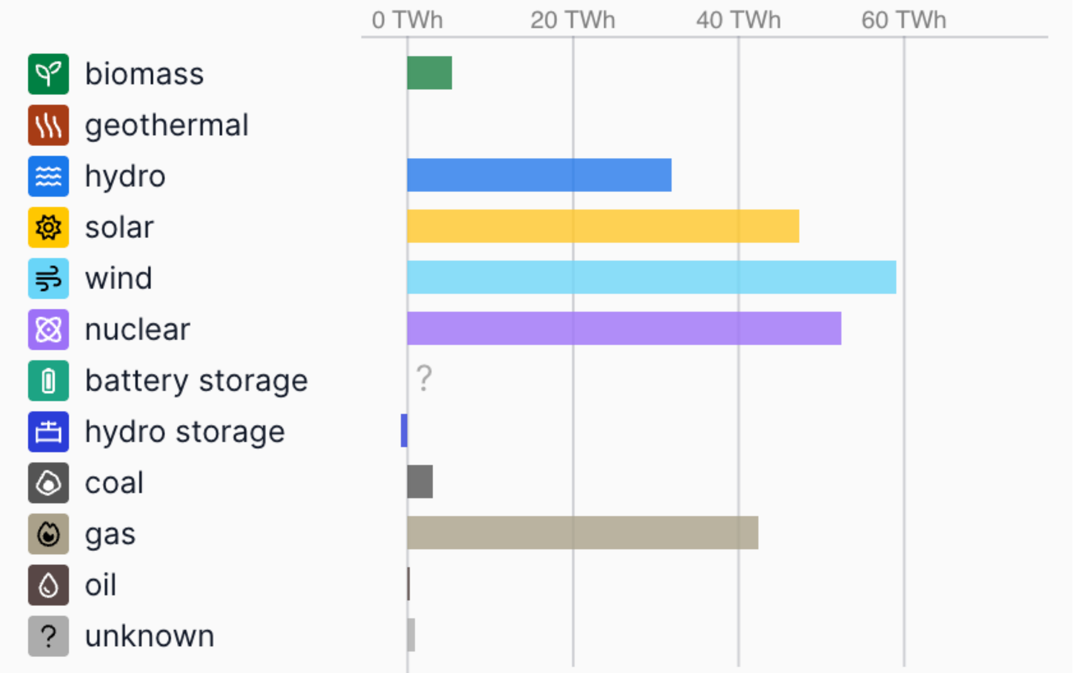




The offsite water footprint is the water usage resulting from using various energy sources to generate electricity for data centers.

Operational Water Footprint (Offsite)





The Energy Water Intensity Factor (EWIF) is dependent on the energy source being used and reflects the intensity of water consumption needed to generate the electricity.

$$H_2O_j^{\text{offsite}} = \text{PUE} \times E_j \times \text{EWIF} \times (1 + \text{WSF}_r^{dc})$$

The Power Usage Effectiveness (PUE) factor measures the energy efficiency of a data center.

WaterWise Water Footprint Modeling of Data Centers

Water Scarcity Factor

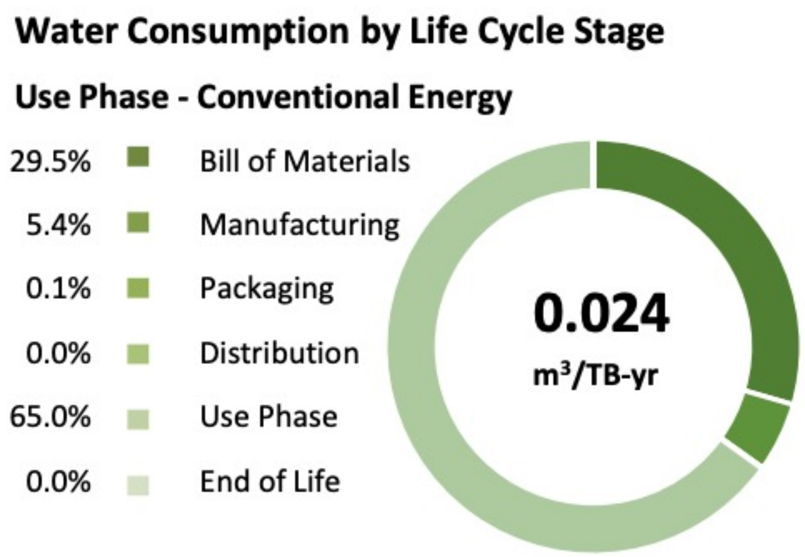
Operational Water Footprint (Onsite)

Operational Water Footprint (Offsite)

Embodied Water Footprint

Embodied Water Footprint

The embodied water footprint is the water consumption generated during server design and manufacturing.



EXOS X16 - 16TB Seagate Sustainability Report

Limited data for CPU, different vendors… WaterWise employs an alternative estimation approach

$$H_2O^{\text{embodied}}_{\text{server}} = E_{\text{manufacturing}} \times \text{EWIF} \times (1 + \text{WSF}^{\text{server}}_{r})$$

Water Footprint: Putting it All Together

$$H_2O_j = H_2O_j^{\text{operational}} + H_2O_j^{\text{embodied}}$$

Water Footprint: Putting it All Together

$$H_2O_j = H_2O_j^{\text{operational}} + H_2O_j^{\text{embodied}}$$

$$H_2O_j = \text{PUE} \times E_j \times \text{EWIF} \times (1 + \text{WSF}_r^{\text{dc}})$$

Onsite Water Footprint

Water Footprint: Putting it All Together

$$\mathrm{H_2O}_j = \mathrm{H_2O}_j^{\text{operational}} + \mathrm{H_2O}_j^{\text{embodied}}$$

$$\mathrm{H_2O}_j = \mathrm{PUE} \times E_j \times \mathrm{EWIF} \times (1 + \mathrm{WSF}_r^{\text{dc}})$$

$$+E_j \times \mathrm{WUE} \times (1 + \mathrm{WSF}_r^{\text{dc}}) + \frac{t_j}{T_{\text{lifetime}}} \cdot \mathrm{H_2O}_{\text{server}}^{\text{embodied}}$$

Water Footprint: Putting it All Together

$$\mathrm{H_2O}_j = \mathrm{H_2O}_j^{\mathrm{operational}} + \mathrm{H_2O}_j^{\mathrm{embodied}}$$

$$\mathrm{H_2O}_j = \mathrm{PUE} \times E_j \times \mathrm{EWIF} \times (1 + \mathrm{WSF}_r^{\mathrm{dc}})$$

$$+E_j \times \mathrm{WUE} \times (1 + \mathrm{WSF}_r^{\mathrm{dc}}) + \frac{t_j}{T_{\mathrm{lifetime}}} \cdot \mathrm{H_2O}_{\mathrm{server}}^{\mathrm{embodied}}$$

Offsite Water Footprint

Water Footprint: Putting it All Together

$$\mathrm{H_2O}_j = \mathrm{H_2O}_j^{\text{operational}} + \mathrm{H_2O}_j^{\text{embodied}}$$

$$\mathrm{H_2O}_j = \mathrm{PUE} \times E_j \times \mathrm{EWIF} \times (1 + \mathrm{WSF}_r^{\text{dc}})$$

$$+E_j \times \mathrm{WUE} \times (1 + \mathrm{WSF}_r^{\text{dc}}) + \frac{t_j}{T_{\text{lifetime}}} \cdot \mathrm{H_2O}_{\text{server}}^{\text{embodied}}$$

Embodied Water Footprint

Water Footprint: Putting it All Together

$$\mathrm{H_2O}_j = \mathrm{H_2O}_j^{\mathrm{operational}} + \mathrm{H_2O}_j^{\mathrm{embodied}}$$

$$\mathrm{H_2O}_j = \mathrm{PUE} \times E_j \times \mathrm{EWIF} \times (1 + \mathrm{WSF}_r^{\mathrm{dc}}) + E_j \times \mathrm{WUE} \times (1 + \mathrm{WSF}_r^{\mathrm{dc}}) + \frac{t_j}{T_{\mathrm{lifetime}}} \cdot \mathrm{H_2O}_{\mathrm{server}}^{\mathrm{embodied}}$$

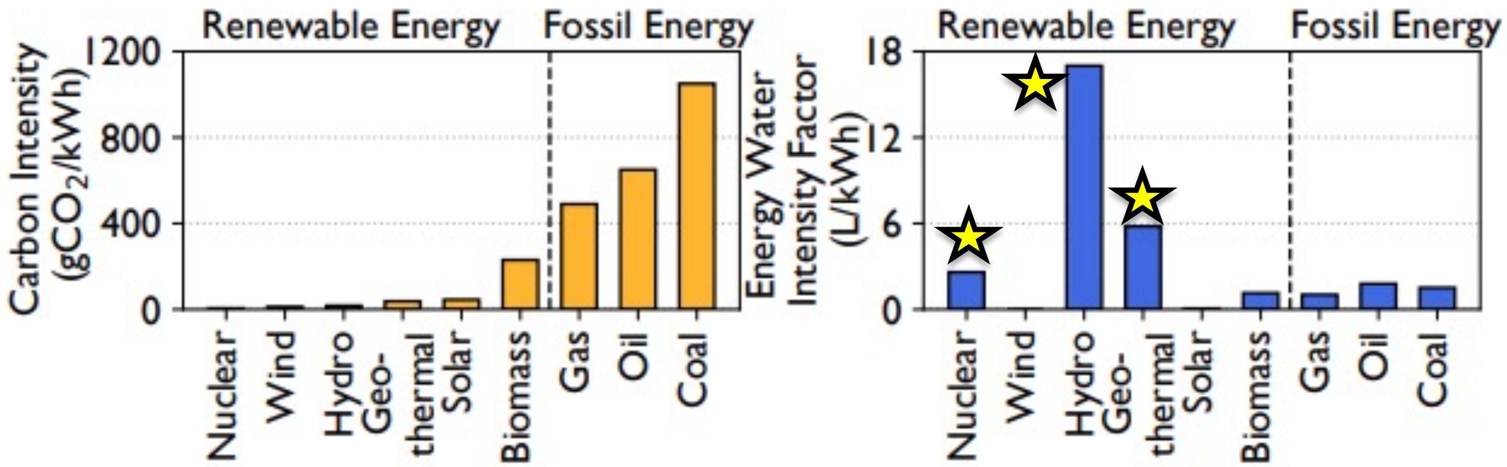
Water Intensity

$$\mathrm{H_2O}^{\mathrm{Intensity}} = (\mathrm{WUE} + \mathrm{PUE} \cdot \mathrm{EWIF}) \cdot (1 + \mathrm{WSF}_r^{\mathrm{dc}})$$

WaterWise: Insights and Key Ideas

… observations, opportunities, and design ideas

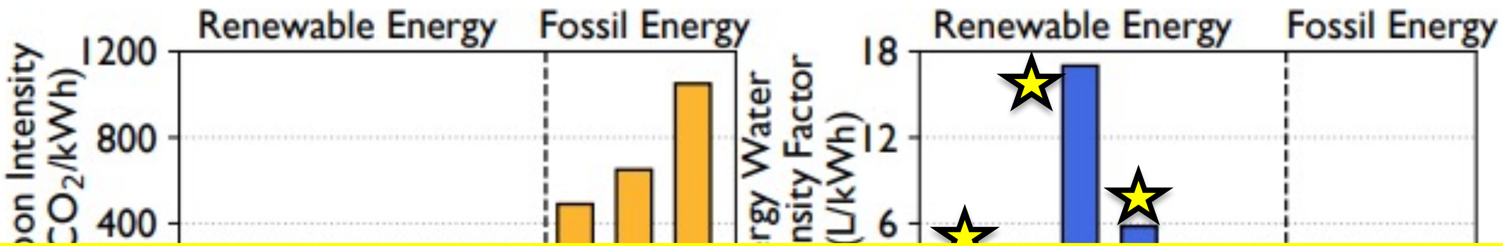
Observation 1: Carbon-friendly energy sources can have higher water footprint requirements.



The amount of water required to generate electricity for powering data centers from different energy sources varies across energy sources.

The water consumption to generate energy from carbon-friendly and carbon-intensive energy sources varies significantly.

Observation 1: Carbon-friendly energy sources can have higher water footprint requirements.

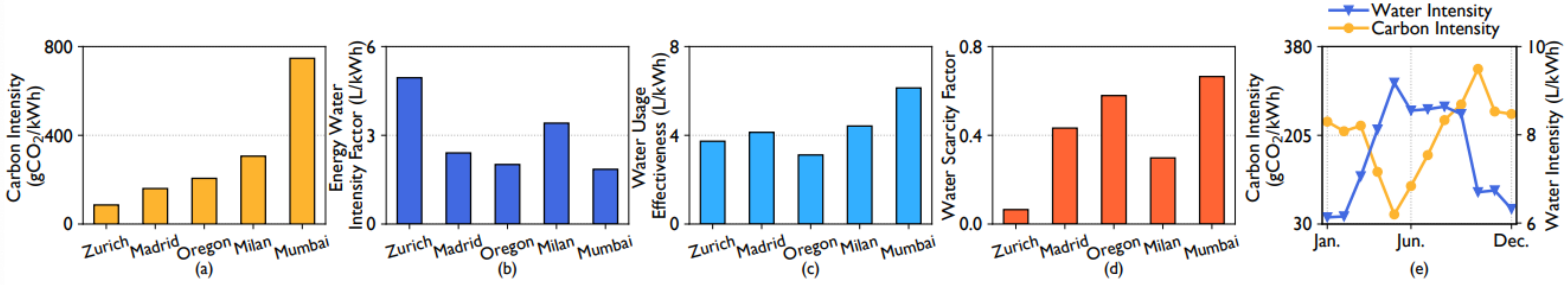


Water and carbon savings are competing goals
Renewable energy sources can increase the water footprint!

The amount of water required to generate electricity for powering data centers from different energy sources varies across energy sources.

The water consumption to generate energy from carbon-friendly and carbon-intensive energy sources varies significantly.

Observation 2: Low Carbon Intensity Regions May be Severely Water-Stressed and High Water Intensity



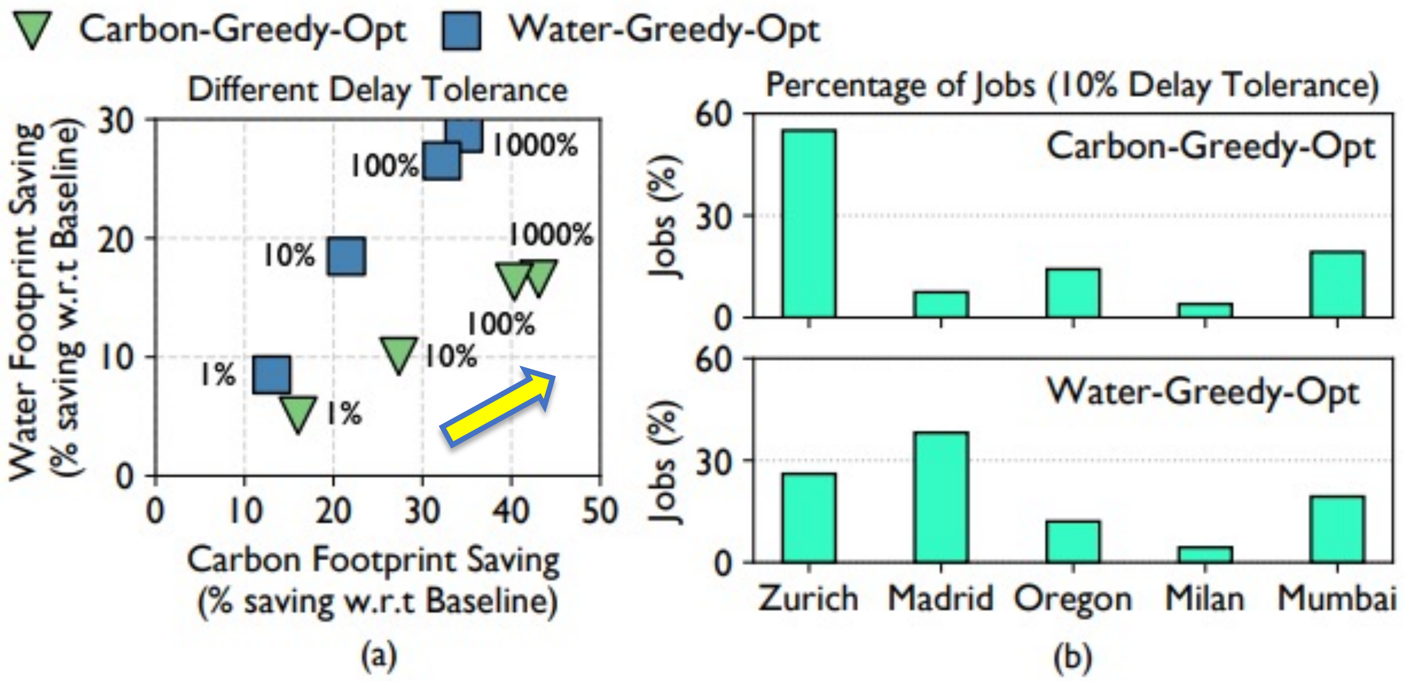
Geographical regions with availability to low carbon intensity energy sources can be water-stressed regions. The carbon-friendly but water-stressed regions may have a relatively higher water scarcity factor.

Consideration of the water scarcity factor changes the trade-off between carbon and water sustainability (carbon vs water intensity).

Key Idea 1: Delay Tolerance Improves the Savings

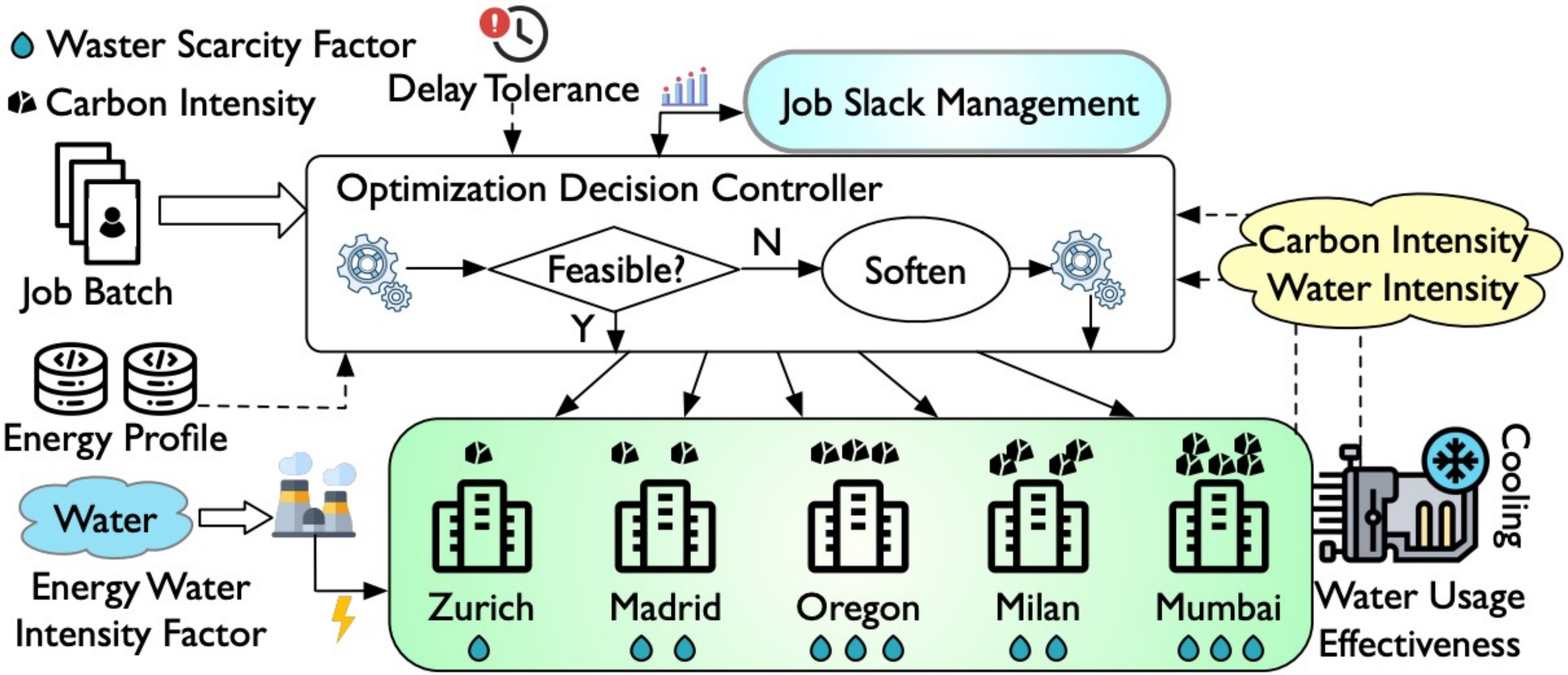
Delay Tolerance: an allowable % increase in the service time of a job compared to its execution time of the job if it had zero transfer latency and queuing delay.

Observation 3: Delay Tolerance Improves the Savings



If jobs could tolerate some latency in their service time – the opportunity for carbon and water footprint savings improves.

WaterWise Optimization Framework



WaterWise Problem Formulation

Formulation

Carbon Part

Water Part

$$\min \sum_{m=1}^{M} \sum_{n=1}^{N} x_{m,n} \cdot \left[\lambda_{CO_2} \frac{CO_2(m,n)}{CO_{2,j}^{max}} + \lambda_{H_2O} \frac{H_2O(m,n)}{H_2O_j^{max}} \right.$$
$$\left. + \lambda_{ref} \cdot \left(\lambda_{CO_2} \cdot CO_{2,n}^{ref} + \lambda_{H_2O} \cdot H_2O_n^{ref}\right)\right]$$

History Part

Key Idea 2: MILP Problem Formulation

Formulation

Carbon Part

Water Part

$$\min \sum_{m=1}^{M} \sum_{n=1}^{N} x_{m,n} \cdot [\lambda_{CO_2} \frac{CO_2(m,n)}{CO_{2,j}^{\max}} + \lambda_{H_2O} \frac{H_2O(m,n)}{H_2O_j^{\max}}$$

$$+\lambda_{\text{ref}} \cdot (\lambda_{CO_2} \cdot CO_{2,n}^{\text{ref}} + \lambda_{H_2O} \cdot H_2O_n^{\text{ref}})]$$

History Part

$$s.t. \quad \sum_{n=1}^{N} x_{m,n} = 1, \quad \forall m \in \{1, 2, \ldots, M\}$$

Assign

$$s.t. \quad \sum_{m=1}^{M} x_{m,n} \leq \text{cap}(n), \quad \forall n \in \{1, 2, \ldots, N\}$$

Capacity

$$s.t. \quad \sum_{n=1}^{N} x_{m,n} \cdot \frac{L_{m,n}}{t_{m,n}} \leq (\text{TOL}\%), \quad \forall m \in \{1, 2, \ldots, M\}$$

Delay Tolerance

The primary objective of WaterWise is to minimize the carbon footprint and water footprint when executing jobs in geographically distributed data centers, under different specified levels of delay tolerance.

WaterWise: Key Ideas

- Key Idea 3: Soft Constraints

When the solver cannot provide feasible and robust optimization outputs, WaterWise softens the delay tolerance constraint.

$$\min \sum_{m=1}^{M} \sum_{n=1}^{N} x_{m,n} \cdot [\lambda_{CO_2} \frac{CO_2(m,n)}{CO_{2,j}^{\max}} + \lambda_{H_2O} \frac{H_2O(m,n)}{H_2O_j^{\max}}$$

$$+\lambda_{ref} \cdot (\lambda_{CO_2} \cdot CO_{2,n}^{ref} + \lambda_{H_2O} \cdot H_2O_n^{ref})] + \sigma \cdot \sum_{m=1}^{M} \sum_{n=1}^{N} P_{m,n}$$

WaterWise: Key Ideas

- Key Idea 3: Soft Constraints

When the solver cannot provide feasible and robust optimization outputs, WaterWise softens the delay tolerance constraint.

$$\min \sum_{m=1}^{M} \sum_{n=1}^{N} x_{m,n} \cdot [\lambda_{CO_2} \frac{CO_2(m,n)}{CO_{2,j}^{\max}} + \lambda_{H_2O} \frac{H_2O(m,n)}{H_2O_j^{\max}}$$

$$+\lambda_{ref} \cdot (\lambda_{CO_2} \cdot CO_{2,n}^{ref} + \lambda_{H_2O} \cdot H_2O_n^{ref})] + \sigma \cdot \sum_{m=1}^{M} \sum_{n=1}^{N} P_{m,n}$$

- Key Idea 4: Job Slack Management

WaterWise designs slack management to determine which jobs are closer to their respective delay tolerance violation (recall that MILP does not retain state information).

$$\text{Urgency} = \text{TOL\%} \cdot t_m - L_m^{\text{avg}} - (T_m^{\text{start}} - T^{\text{current}})$$

Brief Overview of Experimental Methodology

Setup

AWS m5.metal:
eu-central-2 (Zurich)
us-west-2 (Oregon)
eu-south-2 (Spain)
eu-south-1 (Milan)
ap-south-1 (Mumbai)

Metrics

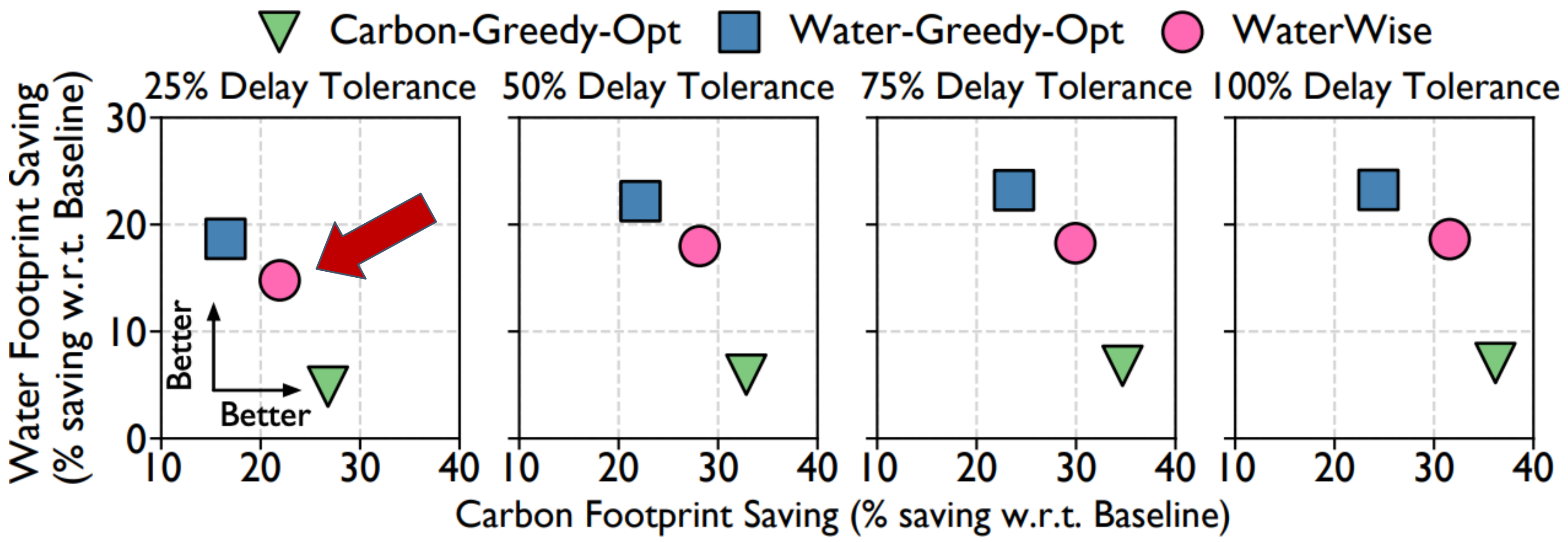
Water Footprint
Carbon Footprint
Latency

Workloads

PARSEC Benchmark
CloudSuite Benchmark

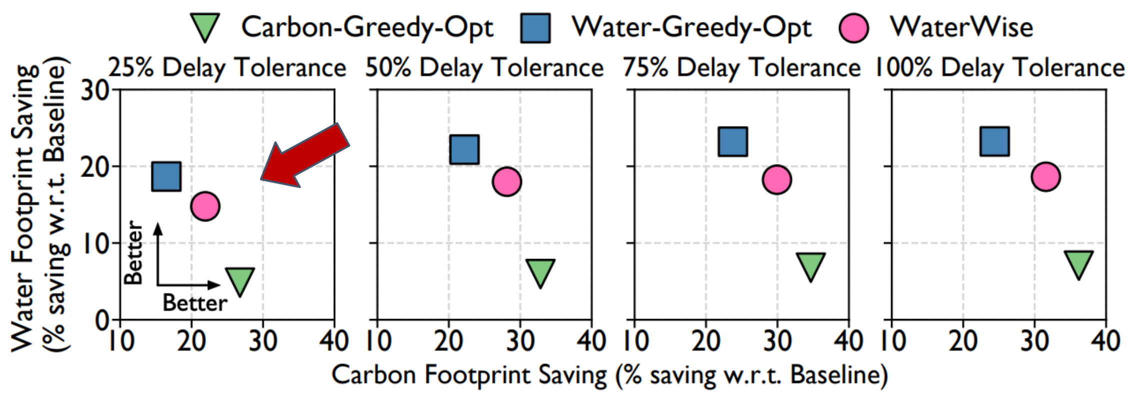
Google Borg trace,
Alibaba Cloud trace

WaterWise Effectively Co-optimizes Carbon and Water Footprint

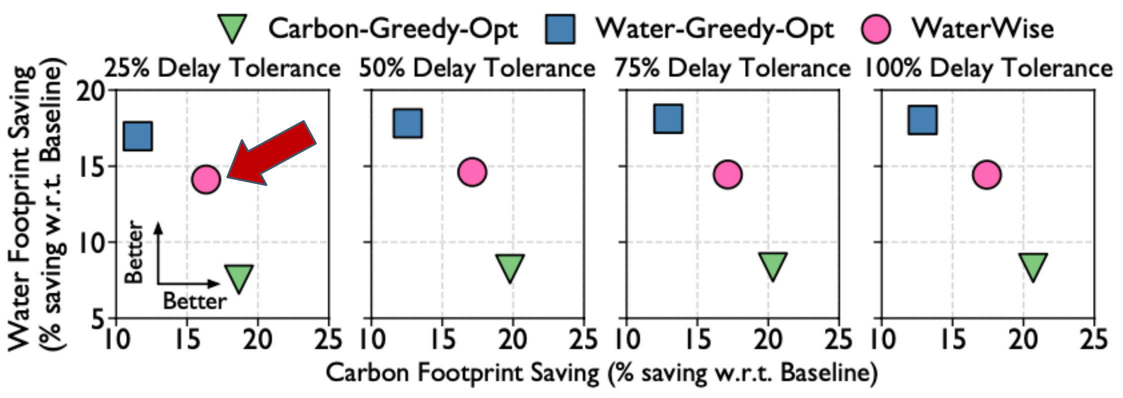


WaterWise provides significant carbon and water footprint savings.
Higher delay tolerance yields additional carbon and water footprint savings.

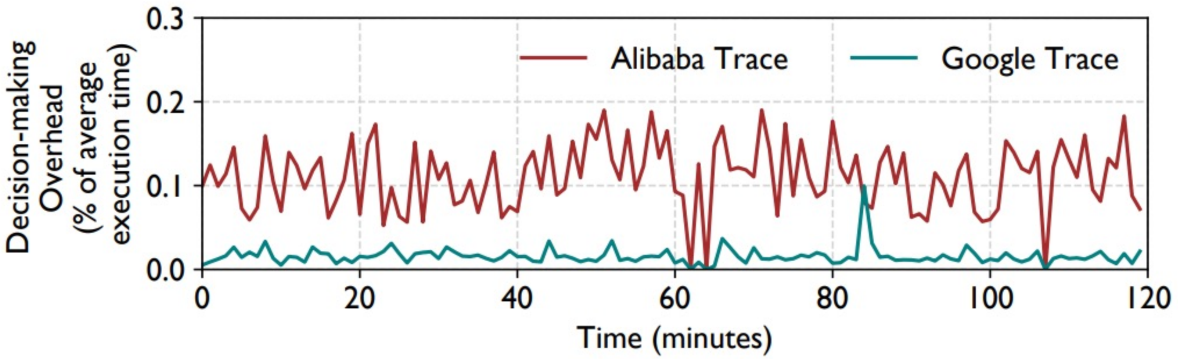
WaterWise Is Practical and Effective Across Google and Alibaba Traces



Google

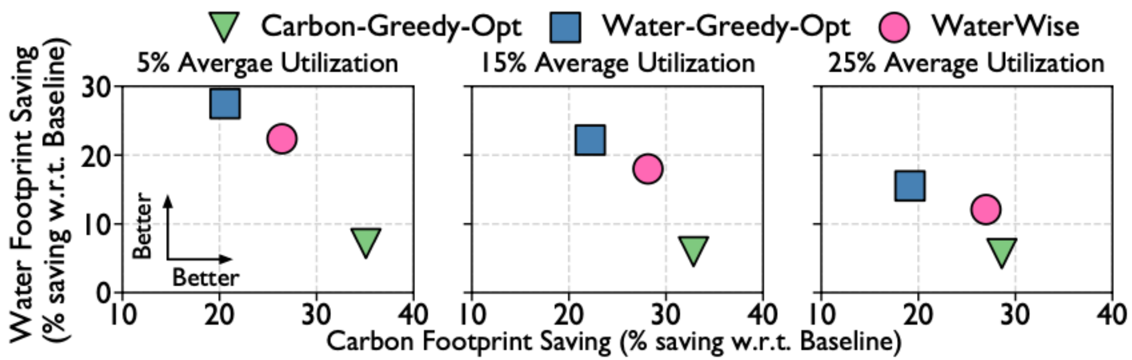


Alibaba

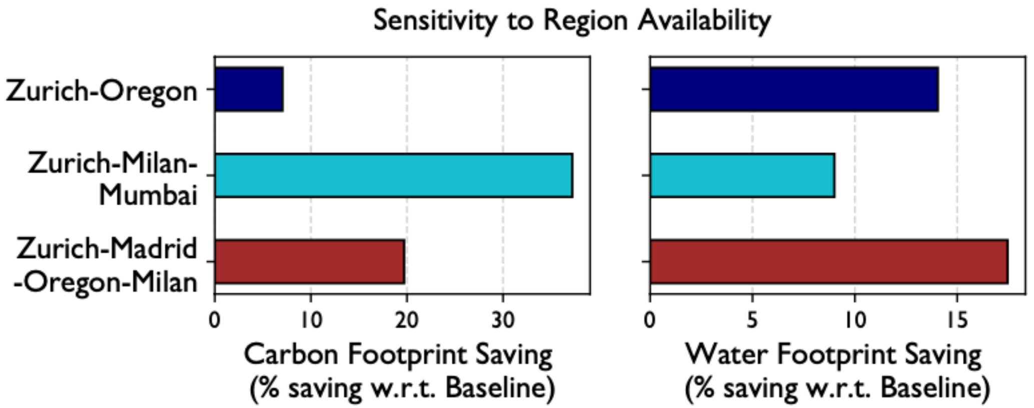


WaterWise is effective across different job arrival patterns and characteristics.

WaterWise Is Effective Across Different Levels of Data Center Utilization and Region Availability



Data center utilization levels



Availability of diverse regions

WaterWise: Summary Key Contributions

- ✓ WaterWise is the first open-source framework to enable exploration of carbon- and water-aware scheduling.
- ✓ WaterWise employs delay tolerance, soft constraints, and slack management to solve the multi-objective optimization.
- ✓ WaterWise provides approx. 20% carbon footprint savings and 14% water footprint savings compared to baseline.

WaterWise

Contact
Yankai Jiang
jiang.yank@northeastern.edu

Bonus Slides

Contact
Yankai Jiang
jiang.yank@northeastern.edu

Sustainability in Computer System
An Emerging Challenge

The Washington Post
Democracy Dies in Darkness

CLIMATE Environment Weather Climate Solutions Climate Lab Green Living Business of Climate

World is on brink of catastrophic warming, U.N. climate change report says

A dangerous climate threshold is near, but 'it does not mean we are doomed' if swift action is taken, scientists say

Clean Energy

Virginia Has the Biggest Data Center Market in the World. Can It Also Decarbonize Its Grid?

Microsoft builds first datacenters with wood to slash carbon emissions

The Carbon Footprint of Amazon, Google, and Facebook Is Growing

How cloud computing—and especially AI—threaten to make climate change worse

Drought-stricken communities push back against data centers

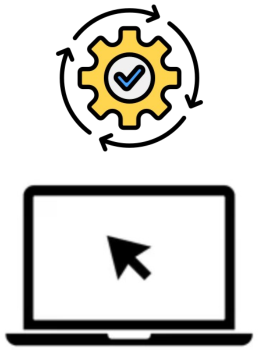
As cash-strapped cities welcome Big Tech to build hundreds of million-dollar data centers in their backyards, critics question the environmental cost.

Microsoft Unveils Zero-Water Datacenter Design to Slash Cooling Water Use

by ESG News • December 12, 2024 Share:

Carbon Emissions Received Significant Attention

Carbon Emissions Received Significant Attention



Operational
Emissions

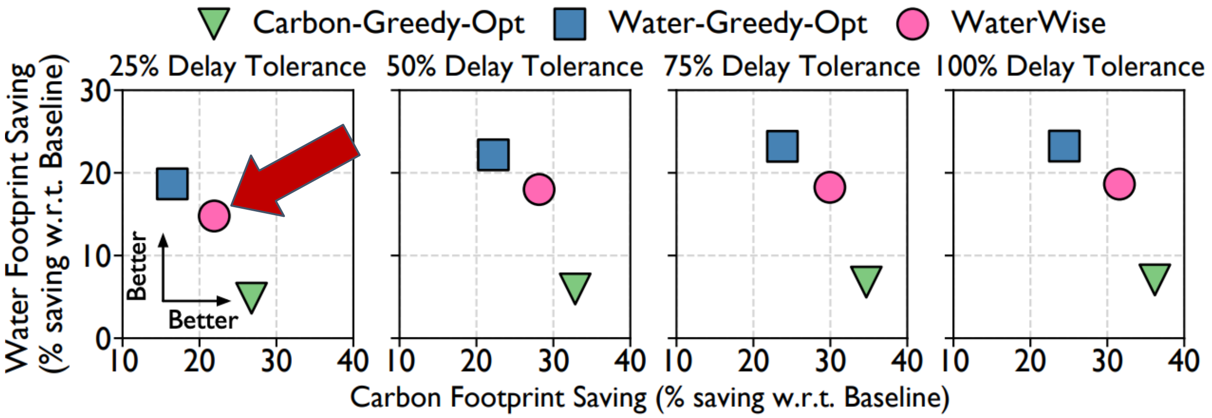
Carbon Emissions Received Significant Attention

Carbon Emissions Received Significant Attention

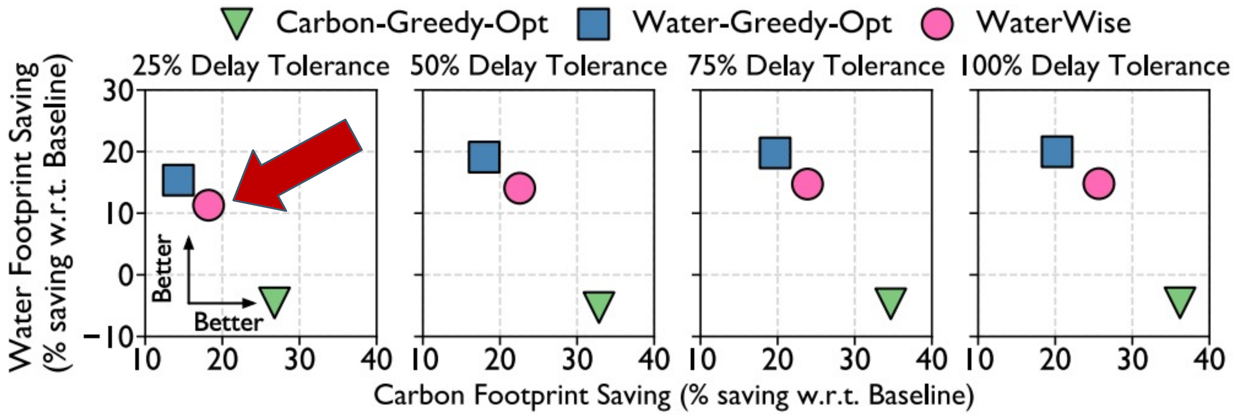
$$CO_{2,j} = CO_{2,j}^{\text{operational}} + CO_{2,j}^{\text{embodied}}$$

$$CO_{2,j} = E_j \cdot CO_2^{\text{Intensity}} + \frac{t_j}{T_{\text{lifetime}}} \cdot CO_{2,\text{server}}^{\text{embodied}}$$

WaterWise is Effective at Reducing Carbon and Water Footprints



Followed the Google Borg trace with 15% utilization



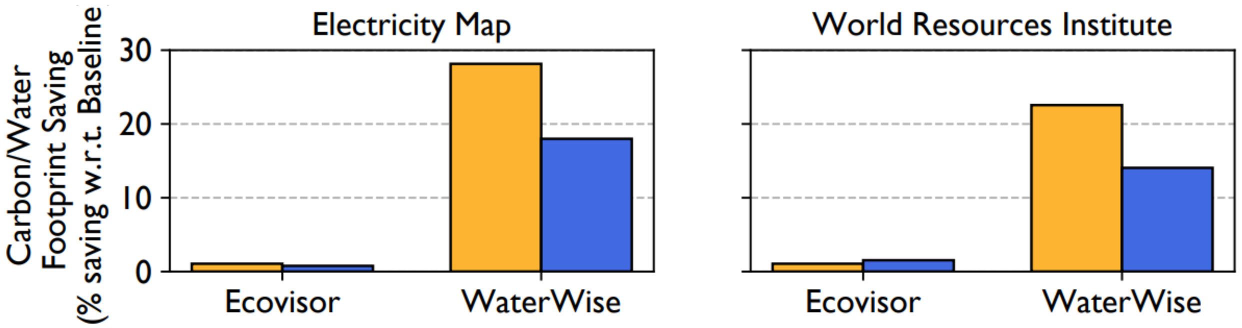
Use different data origin (World Resources Institute)

	Service Time (norm. to execution time)				**% Job Violation**			
Delay Tolerance	25%	50%	75%	100%	25%	50%	75%	100%
Baseline	1×	1×	1×	1×	0%	0%	0%	0%
Carbon-Greedy-Opt	1.06×	1.18×	1.28×	1.50×	1.1%	0.2%	0.3%	0.5%
Water-Greedy-Opt	1.07×	1.17×	1.23×	1.34×	1.7%	0.1%	0.2%	0.2%
WaterWise	1.03×	1.09×	1.11×	1.13×	2.5%	0.03%	0.03%	0.02%

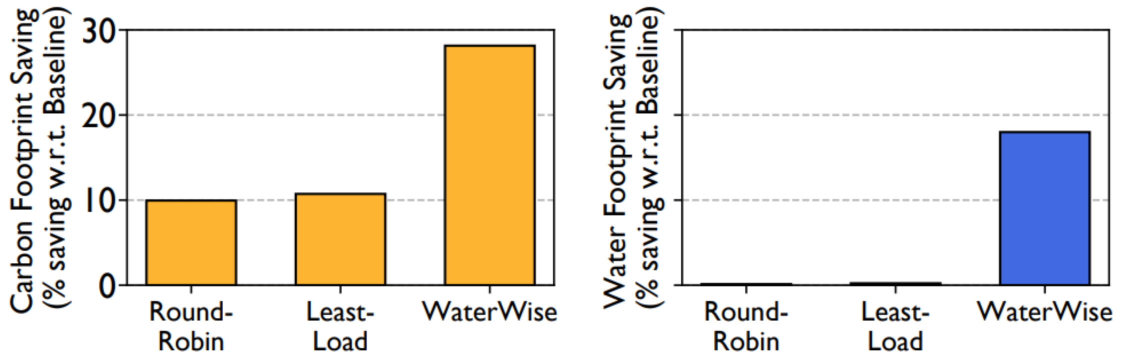
Average service time and delay tolerance violations

WaterWise is effective with low service time and low job Violations. Average service time is much lower than the delay tolerance.

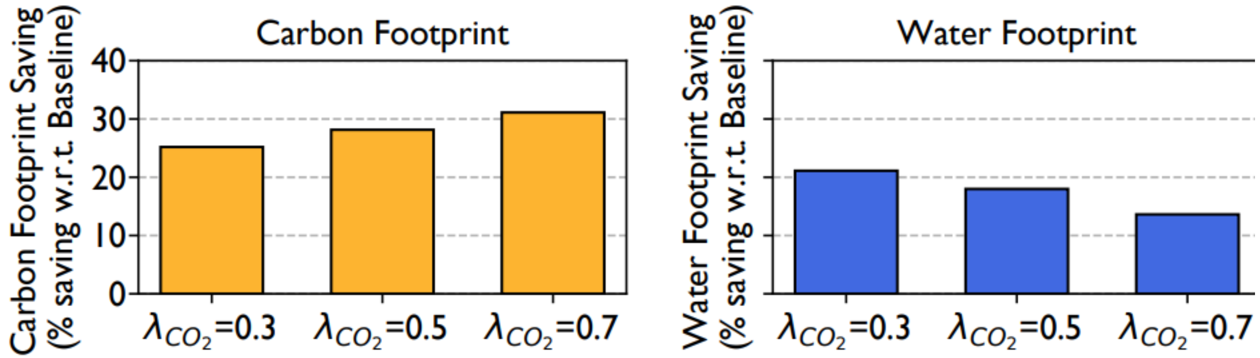
WaterWise is Effective at Reducing Carbon and Water Footprints



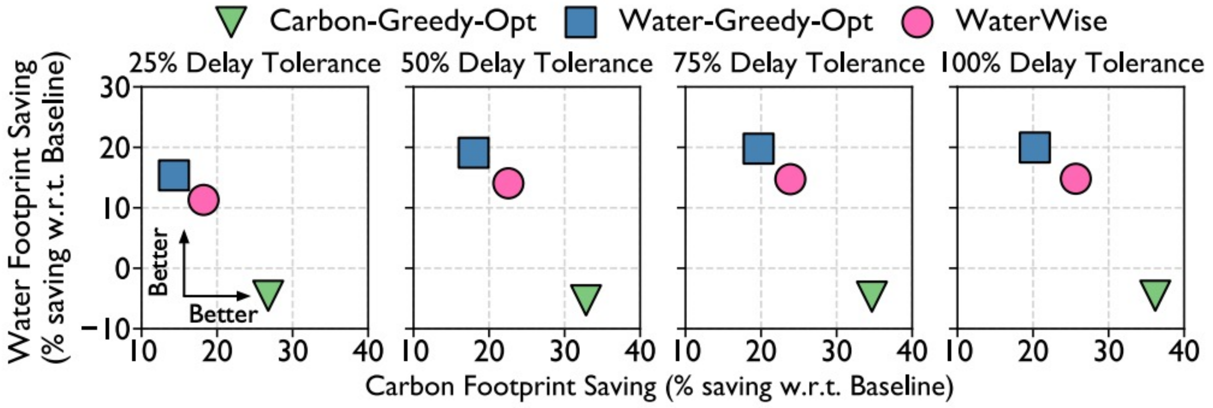
Compared with Ecovisor (ASPLOS' 23)



Compared with alternatives

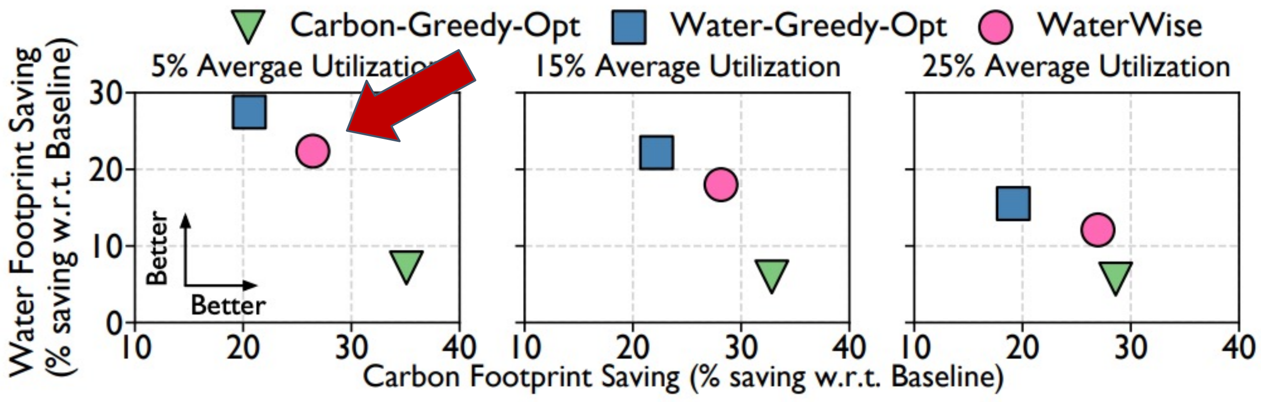


Using different optimization weight factor

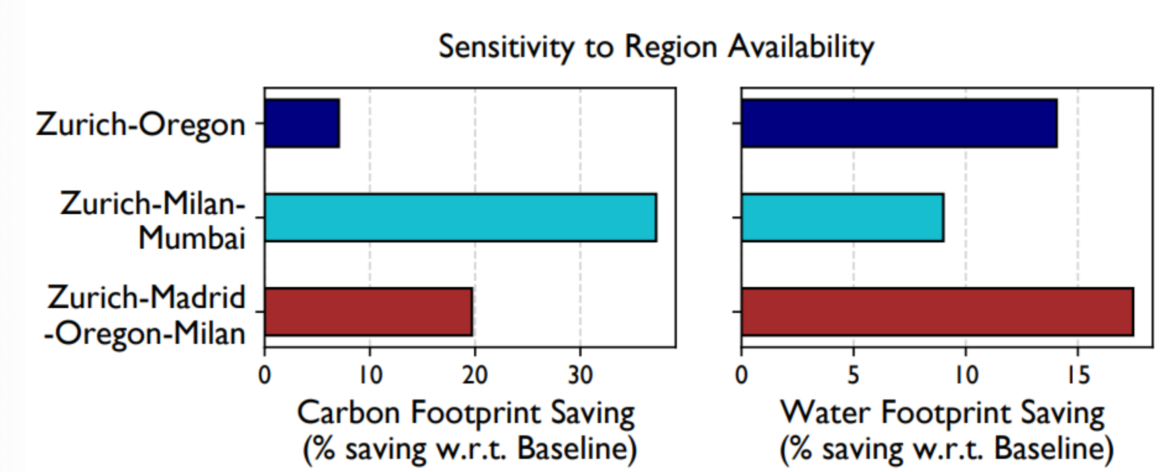


Followed the Alibaba VM traces

WaterWise is Effective at Reducing Carbon and Water Footprints



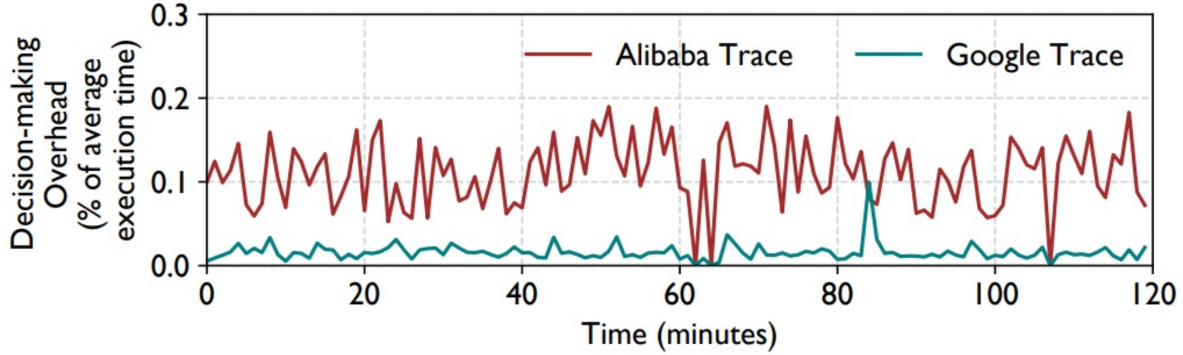
With different data center utilization



With different available regions

Regions	Avg Carbon Overhead (% Execution Carbon)	Avg Water Overhead (% Execution Water)
Zurich	0.08	0.09
Madrid	0.08	0.09
Milan	0.13	0.09
Mumbai	0.17	0.13

Communication overhead (the home region is Oregon)



Low decision-making overhead