Carbon Emissions from Cloud Storage:
An Empirical Analysis toward Actionable User Practices for Sustainable Storage

Gopika Premsankar, Lin Wang, Suzan Bayhan

HotCarbon 2026 ▪ July 16, 2026

Aalto University
School of Electrical
Engineering

PADERBORN
UNIVERSITY

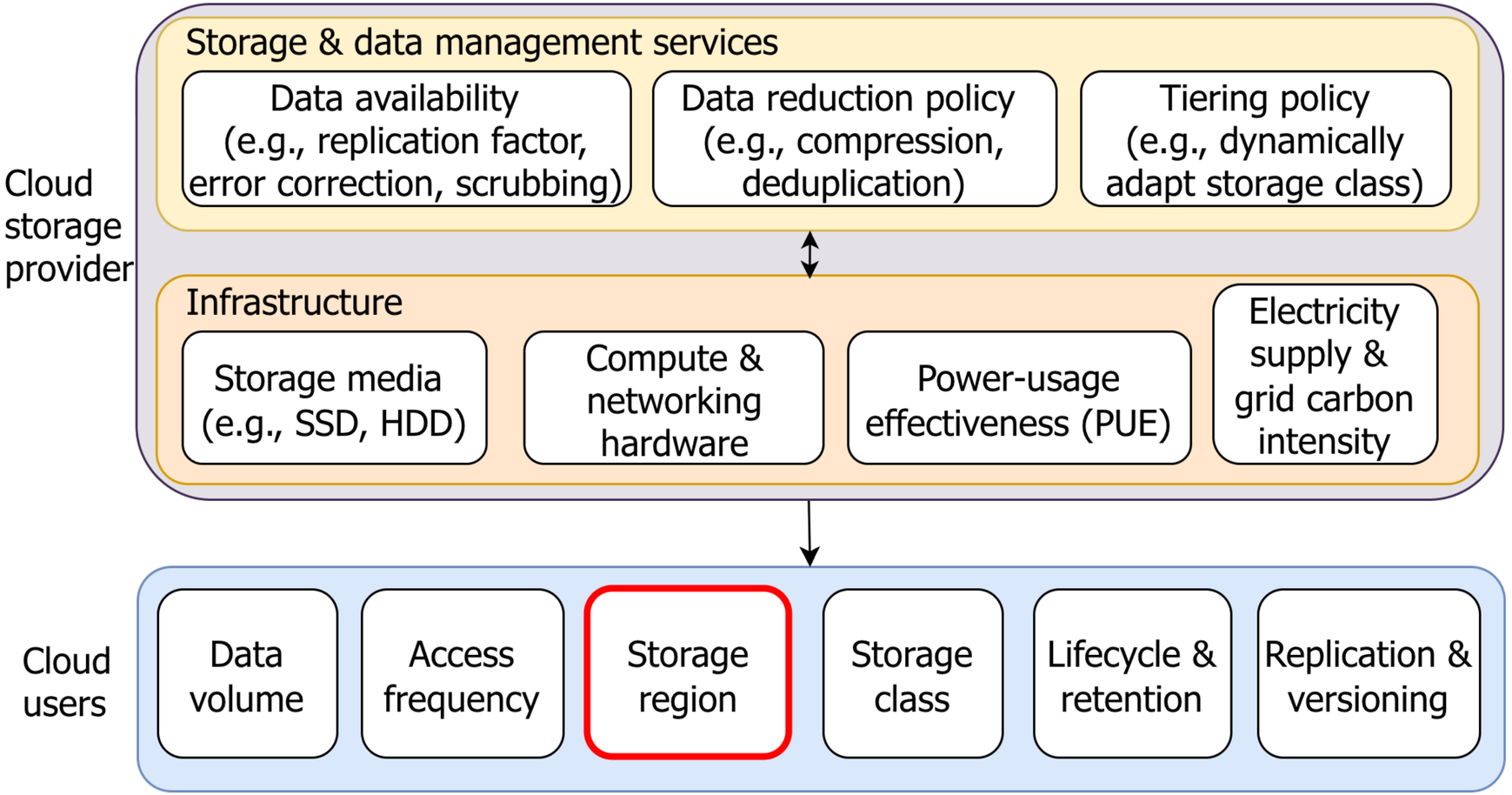
UNIVERSITY
OF TWENTE.

Object storage

- Backbone of many applications today, ranging from social media to backups
- Emissions from storage services are under-studied although they contribute to overall operational emissions: e.g. up to 33% of Azure's operational emissions

Sara McAllister et al. "A call for research on storage emissions." HotCarbon 2024.

Operational emissions from object storage



Sara McAllister et al. "A call for research on storage emissions." HotCarbon 2024.
Varsha Rao and Andrew A. Chien. "Understanding the operational carbon footprint of storage reliability and management." HotCarbon 2024.

Cloud provider recommendations

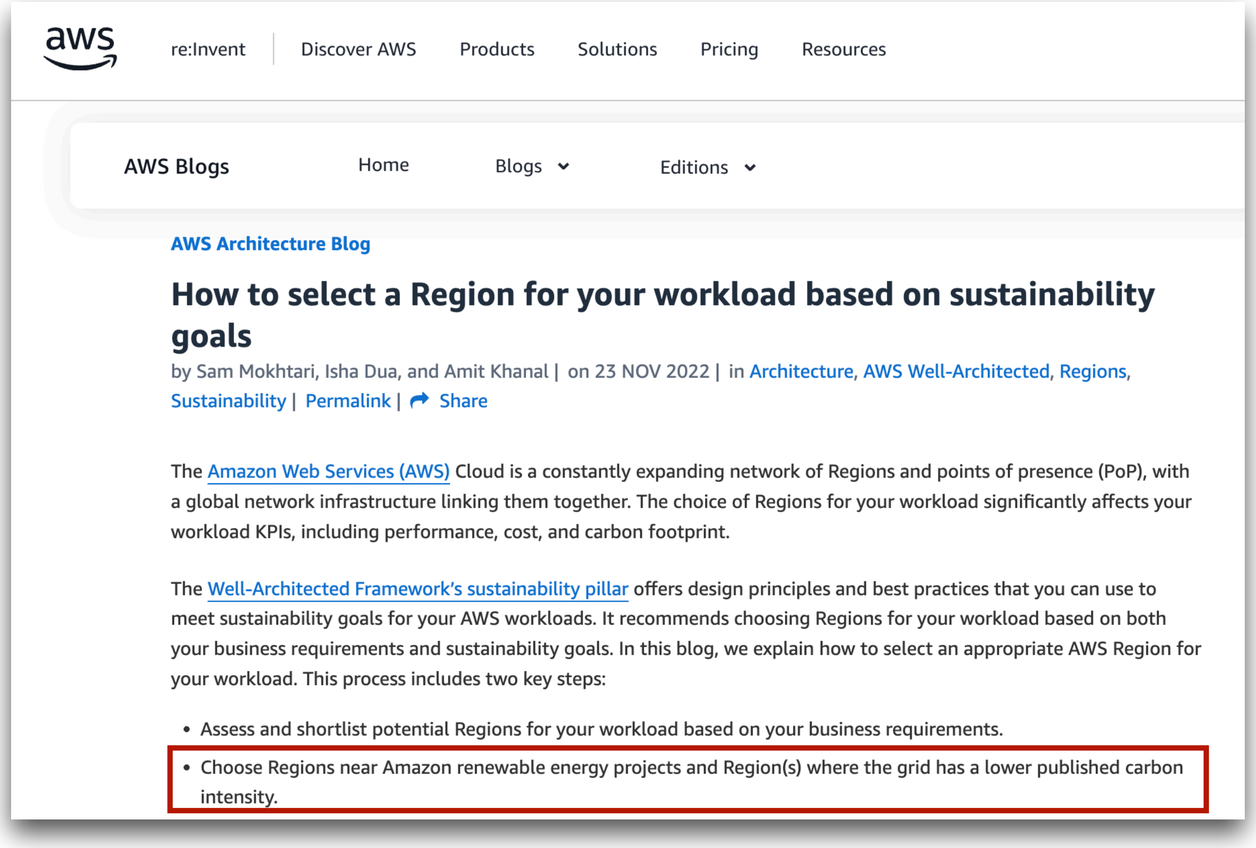
Choice of location to reduce emissions

How to incorporate carbon free energy in your location strategy

Be sure to consider the other best practices for choosing resource locations like data residency requirements, latency to your end users, redundancy of the application, and price of the services available.

To use the CFE data above, here are some good ideas to get you started:

1. **Pick a cleaner region for your new applications.** If you are going to run an application over time, running in the region with the highest CFE% will emit the lowest carbon emissions.
2. **Run batch jobs on the cleanest option.** Batch workloads often have the benefit of planning. You should pick the region with the highest CFE% available to you.
3. **Set an organizational policy for low carbon locations.** You can restrict the location of your resources to a particular Google Cloud region or subset of regions using the "Resource Location Restriction" organization policy. Dedicated "low carbon" value groups have been created to enable you to restrict locations with low carbon impact. For example, if you want to use only US-based regions, use the "Low carbon United States" (`in:us-low-carbon-locations`) value group.



https://cloud.google.com/sustainability/region-carbon

https://aws.amazon.com/blogs/architecture/how-to-select-a-region-for-your-workload-based-on-sustainability-goals/

Motivation

How does choosing the storage region affect emissions from data hosted in cloud object storage?

Experimental setup

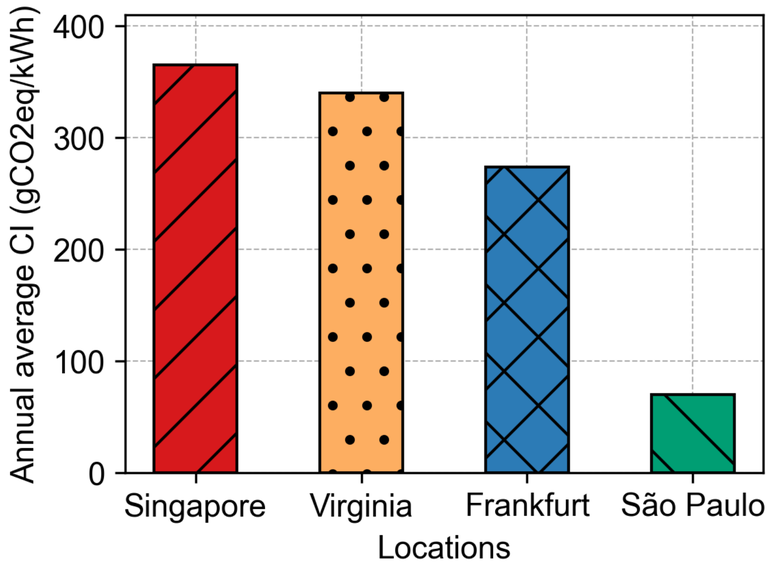
Store 1TB data in four regions of AWS S3 and GCS between Jun—Nov 2025

→

Retrieve monthly location-based emissions reports

→

Analyse data w.r.t public CI and PUE data



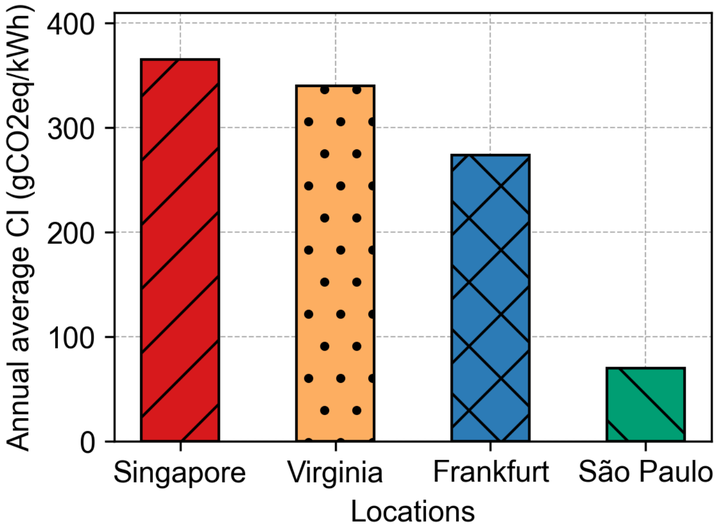
- AWS Customer Carbon Footprint Tool
- Google Cloud Carbon Footprint

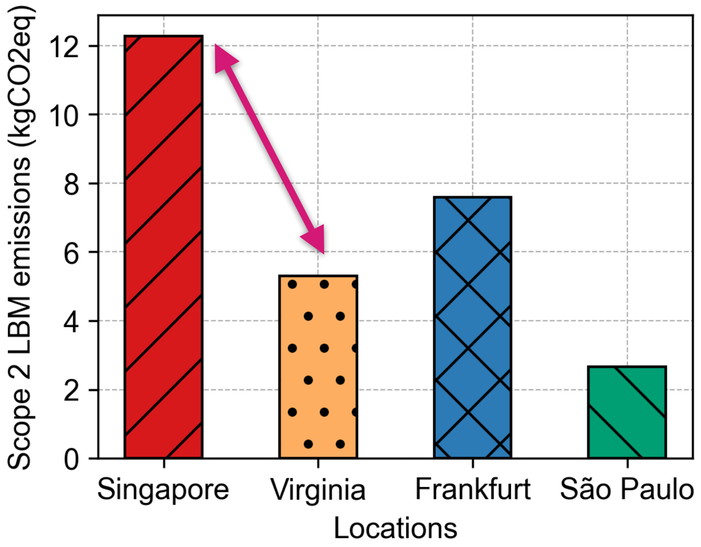
ELECTRICITY MAPS

Experimental results

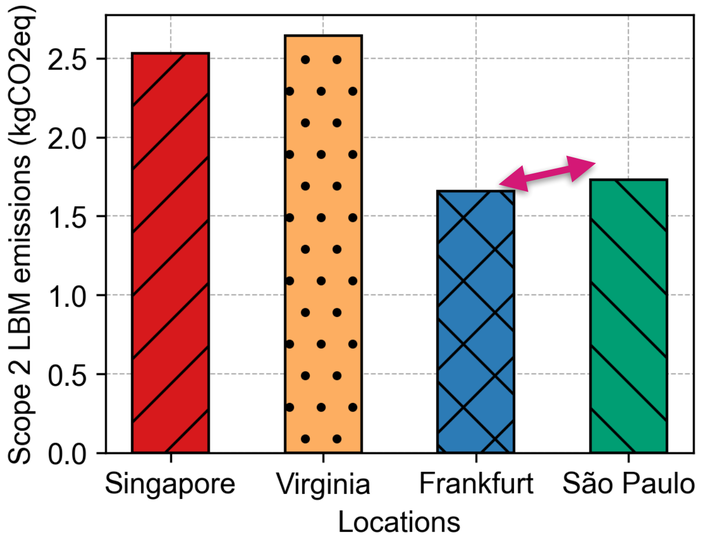
- RQ1: Do locations with lower CI result in lower carbon emissions reported for data stored in those regions?
- RQ2: How does the electricity used in each location vary?

RQ1: Reported carbon emissions over 6 months





AWS



GCP

- Given the CI of the regions, expected: lowest emissions in São Paulo and similarly high emissions in Singapore and Virginia
- Unexpected trends in the results

Takeaway #1

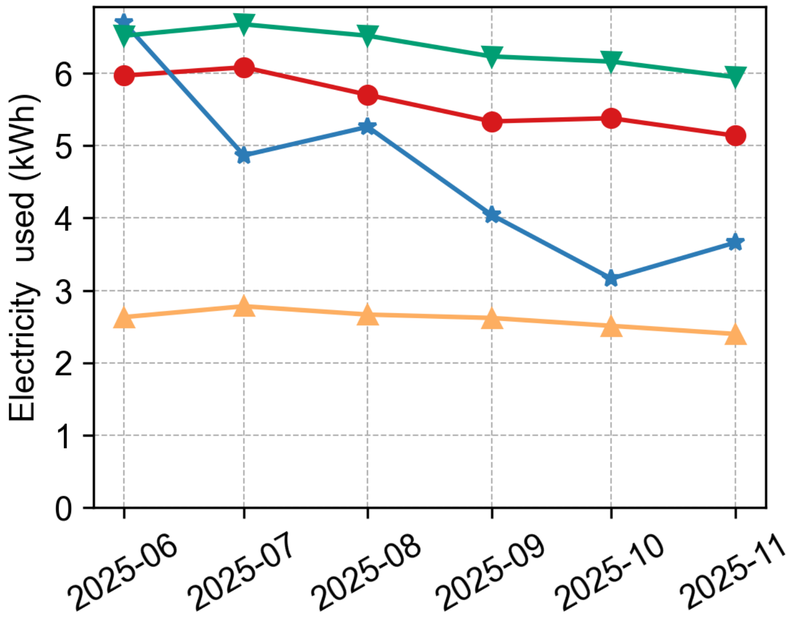
Choosing a region purely based on lower CI does not necessarily result in lower reported location-based emissions.

Research questions

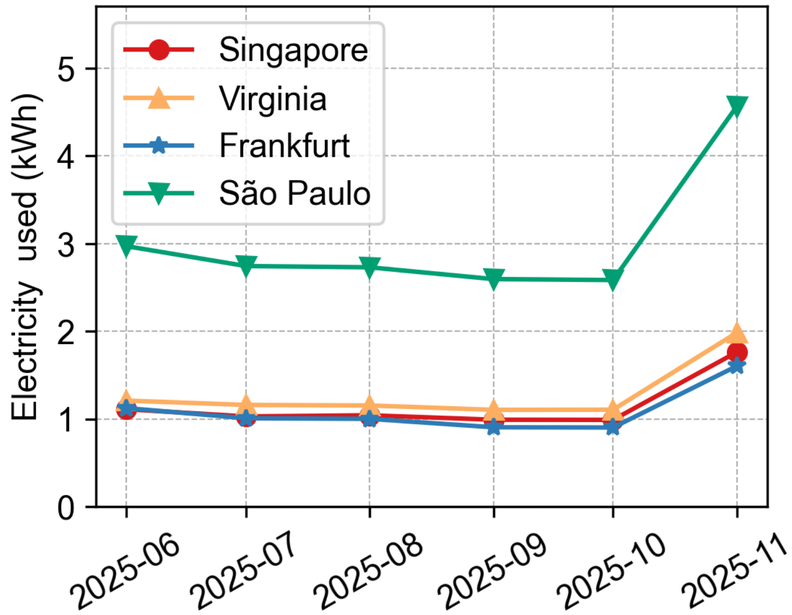
- RQ1: Do locations with lower CI result in lower carbon emissions reported for data stored in those regions?
- RQ2: How does the electricity used in each location vary?

RQ2: Electricity usage

$$E_{\text{storage,loc}} = \frac{C_{\text{storage,loc}}}{CI_{\text{loc}}}$$



AWS



GCP

https://docs.aws.amazon.com/sustainability/latest/userguide/energy-calculation.html

Takeaway #2

Electricity consumption is highest in the region with the lowest CI, indicating that its IT equipment and/or facilities are less energy-efficient than those in other regions.

Future research

- A carbon score that incorporates both carbon-free energy (CFE) and CI, in addition to other proxies for electricity use

$$S_{loc} = \left(1 - \frac{CFE_{loc}\,\%}{100}\right) \times CI_{loc} \times PUE_{loc}$$

- Optimize storage locations based on new score; consider additional constraints / levers such as cost, data locality, storage class

Thank you

Contact:

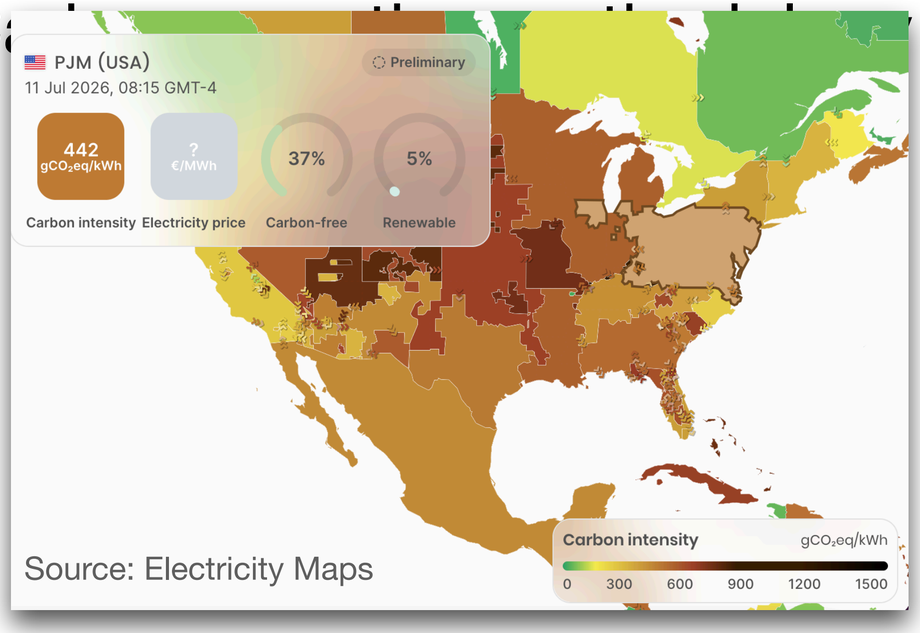
- gopika.premsankar@aalto.fi
- lin.wang@uni-paderborn.de
- suzan.bayhan@utwente.nl

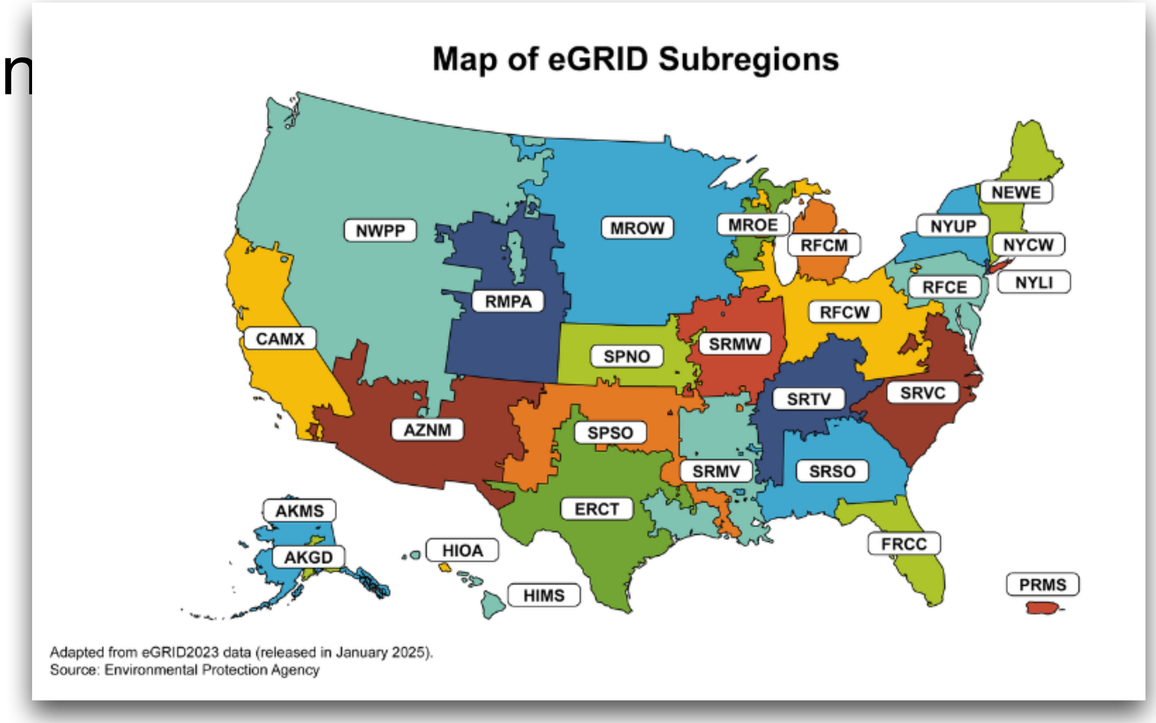
Acknowledgments: Sujay Srivastava and Mehrdad Aslani for experiment setup and data collection.

Extra slides

Limitations

- Choice of carbon intensity data
- C[illegible] chan





Source of eGRID map:
https://www.buildingenclosureonline.com/blogs/14-the-be-blog-building-enclosure/post/93764-using-egrid-data-to-approximate-operational-carbon

Carbon reporting methodology

$$C_{\text{storage,loc}} = E_{\text{storage,loc}} \times CI_{\text{loc}}$$

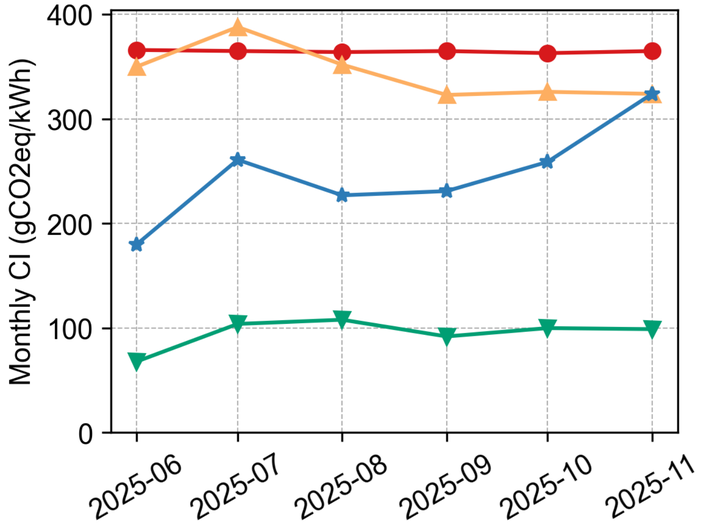
- Estimated (or measured) electricity used by cluster; allocation methods to the customers could differ
- Overhead of ancillary services such as cooling and lighting considered by Google

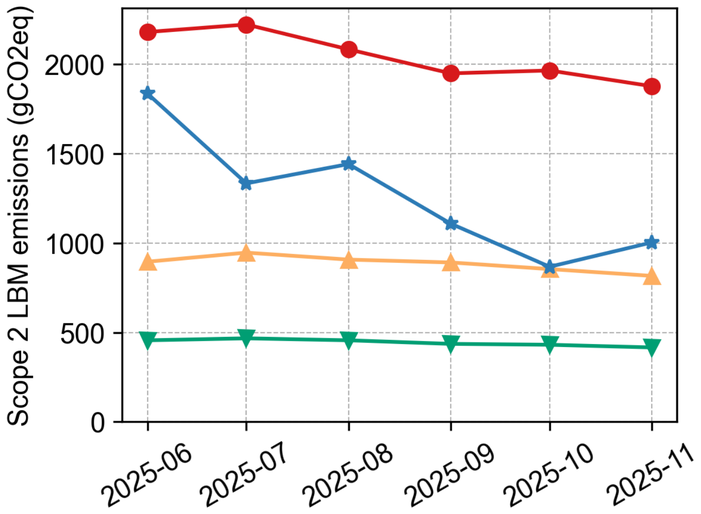
Carbon reporting methodology

$$C_{\text{storage,loc}} = E_{\text{storage,loc}} \times CI_{\text{loc}}$$

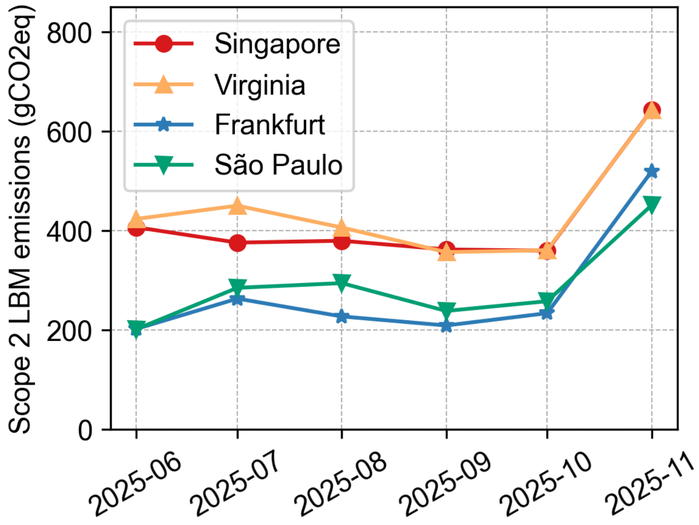
- Google sources hourly CI data from Electricity Maps
- AWS relies on supplier-provided information, and other sources: eGRID in the U.S., IEA; time resolution unknown

Carbon emissions over time





AWS



GCP