Carbon-aware inter-datacenter workload scheduling and placement

UCSD CSE Computer Science and Engineering

Yibo Guo, George Porter

E-lab

Background

Data centers are consuming more energy and thus producing more carbon emissions. To deal with the growing carbon footprint of computing, operators have begun adopting low-carbon renewable energy sources like solar and wind.

However, this is not a trivial task, as these renewables often vary across time of day or geographical locations. This creates an interesting and non-trivial optimization problem where:

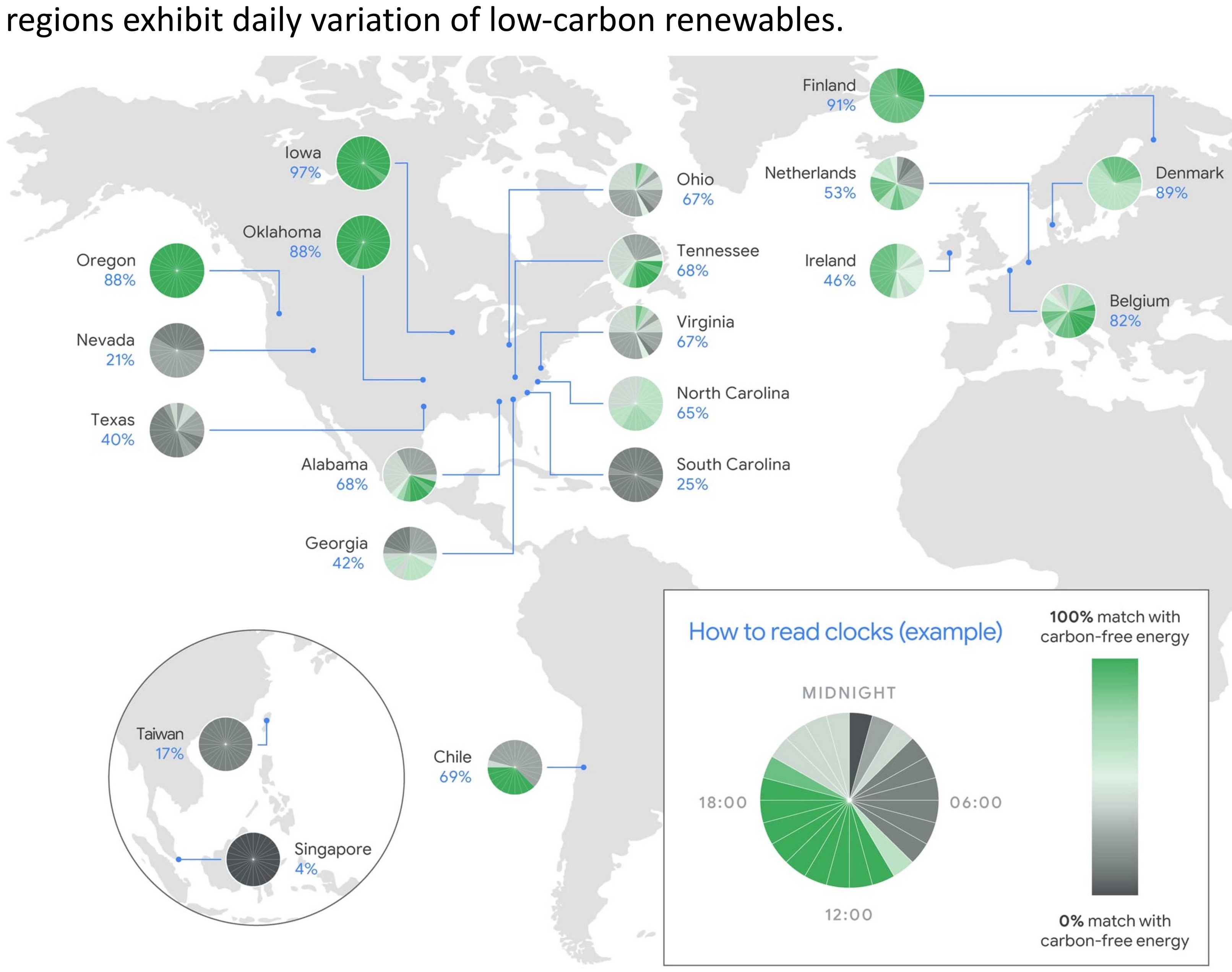
- a *carbon-agnostic* approach runs workloads when/where they arrive, but
- a *carbon-aware* approach moves workloads in time or space to take advantage of available low-carbon renewable energy.

Existing research focuses on *time-shifting* workloads by delaying work from high-carbon hours to low-carbon hours, but shifting within a single site limits the carbon saving potentials. In this work, we investigate space-shifting, a more involved process that spans across multiple sites. Although space-shifting may result in higher overhead, it has the potential to generate even greater carbon savings.

Intermittent renewable energy availability

The availability of renewables can fluctuate significantly depending on the time of day and geographical location, owning to their inherent characteristics. For example, solar peaks around mid-day and hydro/wind are highly regional.

The figure below shows the renewable availability in Google's data centers as of 2022. Note that some regions are cleaner (greener) than other regions and most regions exhibit daily variation of low-carbon renewables.



Renewable availability in Google's Data Centers in 2022 (Source: cloud.google.com)
Low-carbon power (in green) varies across both time and space

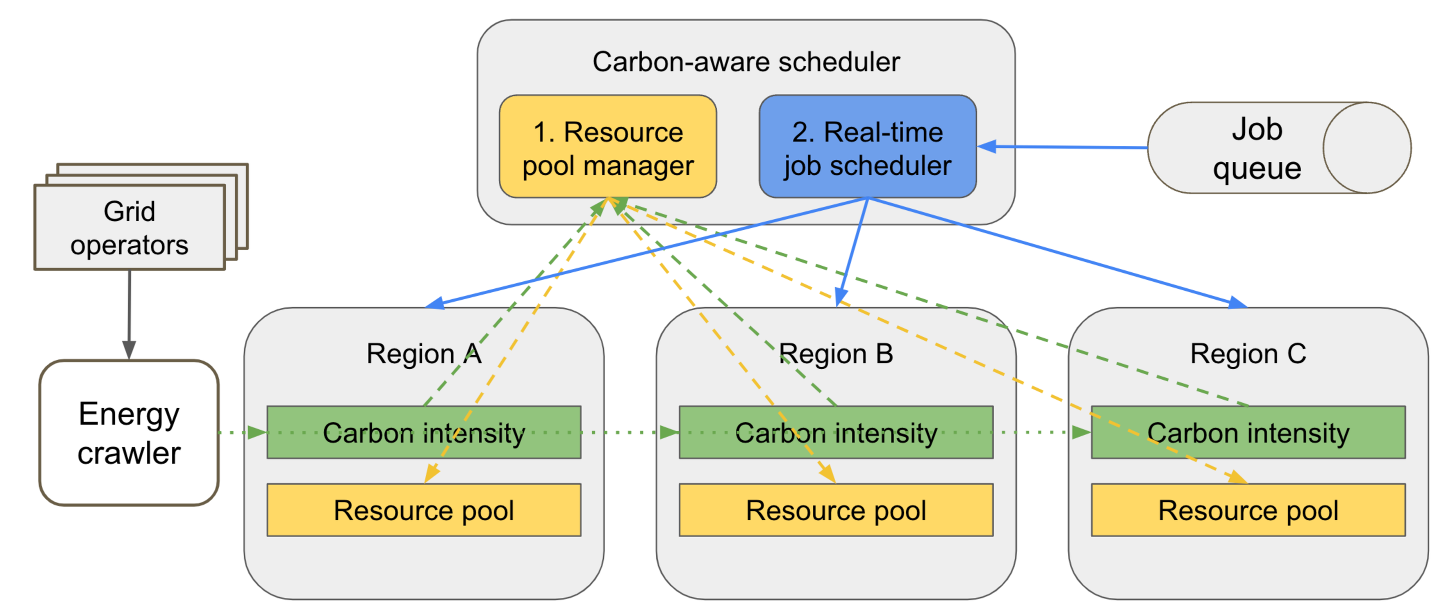
Learn more on our website:

Proposed space-shifting solution

We propose a geo-distributed cloud scheduling platform that maximizes the use of low-carbon power.

At a high level, we employ a two-level scheduling system that:

1. adjusts resource footprint across regions based on carbon cleanness,
2. assigns individual jobs to their optimal locations while considering both carbon savings and migration cost.



The resource pool manager adjusts the available resources in each region, based on the carbon cleanness and current utilization:

1. if there are more workloads, it prioritizes the regions with low-carbon power;
2. if there are fewer workloads, it reduces the resource footprint in high-carbon regions. This operates at a lower frequency, e.g. 15min.

The job scheduler makes real-time decisions on where to run a job, by considering:

1. available resources: based on capacity at each region,
2. carbon cleanness: from crawled energy data, and
3. migration cost: based on data size and WAN bandwidth.

The goal is to avoid moving jobs with high migration cost that negates the carbon savings.

Carbon savings vs migration cost

To balance between carbon savings and migration cost, we calculate:

- the overhead in terms of additional energy consumption of moving a workload (+X%), and compare it with
- the carbon savings of such movement (−Y%).

If X is on par with Y, then it's not worthwhile to move this workload, but if X << Y, this means that we can achieve carbon savings with relatively negligible overhead.

This intuition incentivizes us to define this new metric to guide migration decisions, which is the ratio of a job's predicted compute energy usage and its input/output data size. More formally, we define:

$$\texttt{migration cost index} = \frac{\textit{Predicted compute energy usage}}{\textit{Predicted data size}}$$

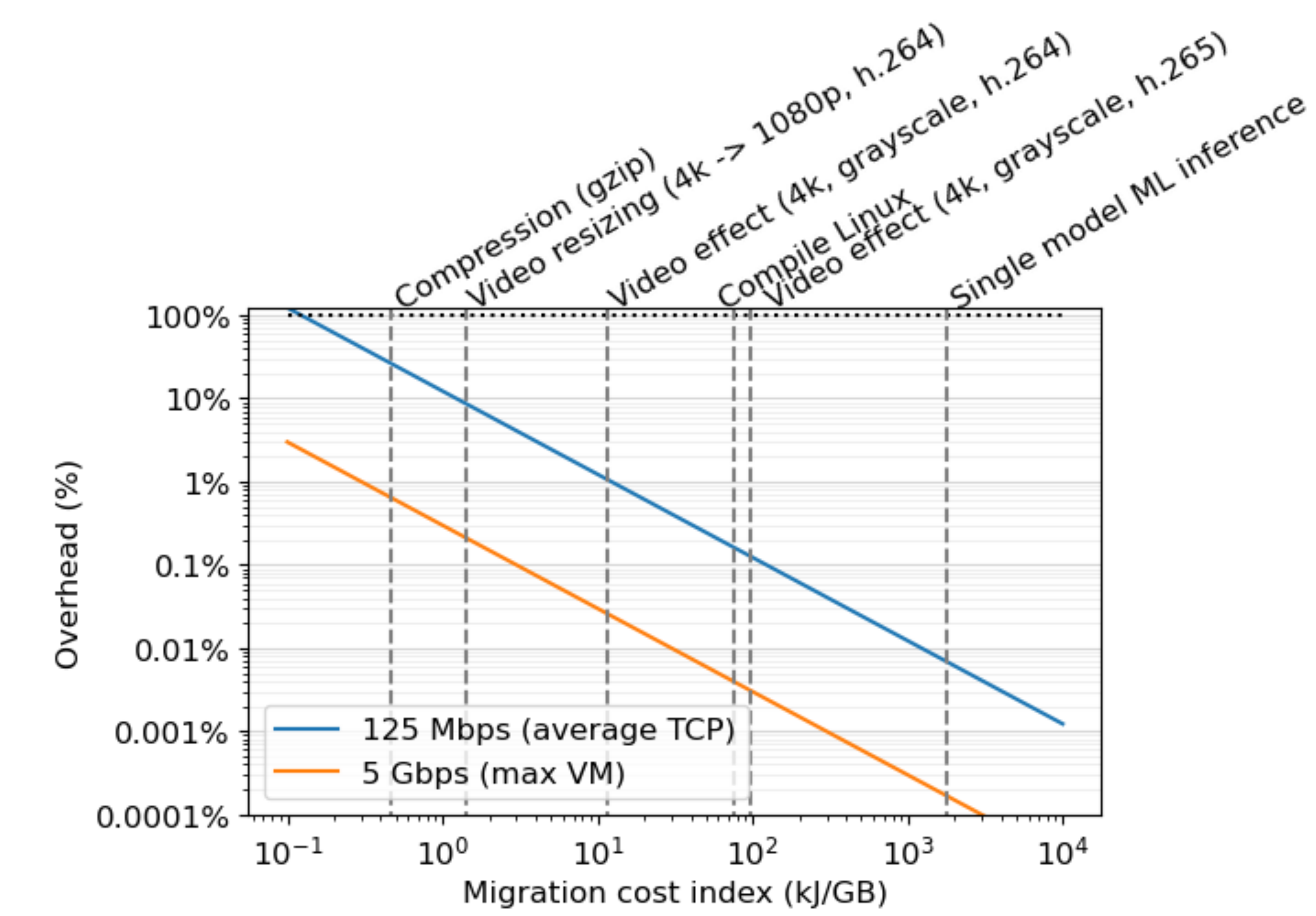
We then use this index to guide our migration decisions among a large set of jobs.

Profiling migration cost index

We profiled a set of candidate workloads on semi-modern servers with Intel Xeon CPUs, and graphed the overhead under different bandwidth constraints, using wide-area VM bandwidth measured in previous studies.

Workload	Migration cost index (kJ/GB)
Compression (gzip)	0.47
Video resizing (4k → 1080p, h.264)	1.41
Video effect (4k, grayscale, h.264)	11.53
Compile Linux	76.42
Video effect (4k, grayscale, h.265)	96.8
Single model ML inference	1800

Overhead analysis on candidate workloads



Research questions

- How to accurately predict workload energy usage and dataset size?
- What happens when input datasets are already geo-replicated?
- What if multiple jobs depend on a few shared datasets?
- How to deal with renewable variation with fixed compute capacity?
- How to effectively reduce total power usage at high-carbon locations?

Evaluation plan

We are currently implementing our system on Nautilus, a Kubernetes research platform that supports three US regions.

We are building a `kubectl` wrapper so users can seamlessly opt in. We plan to provide insights like predicted carbon savings, estimated migration cost and actual net savings.