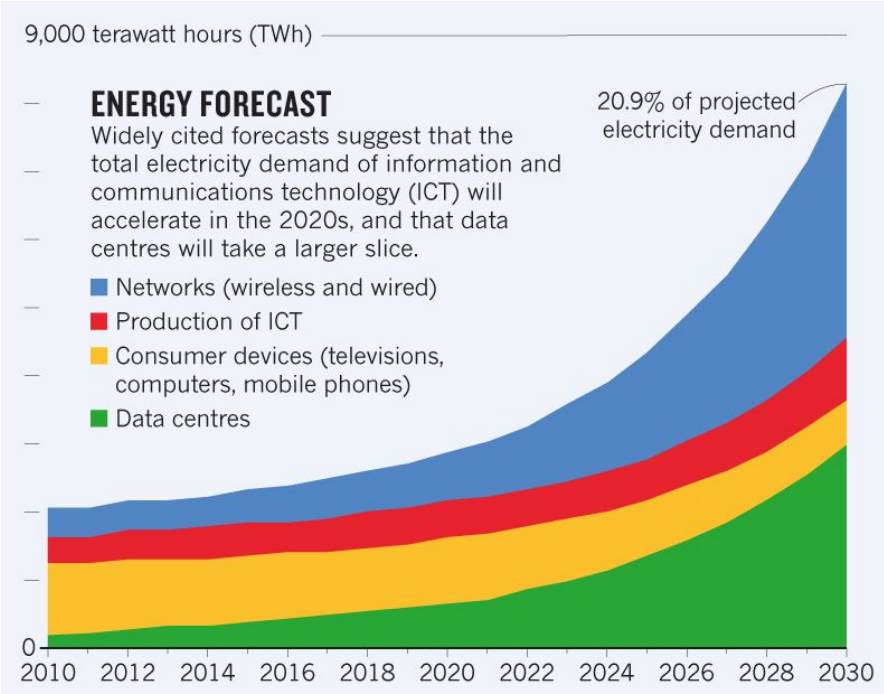
A metric for factoring data movement into chasing the sun

Yibo Guo, George Porter
UC San Diego

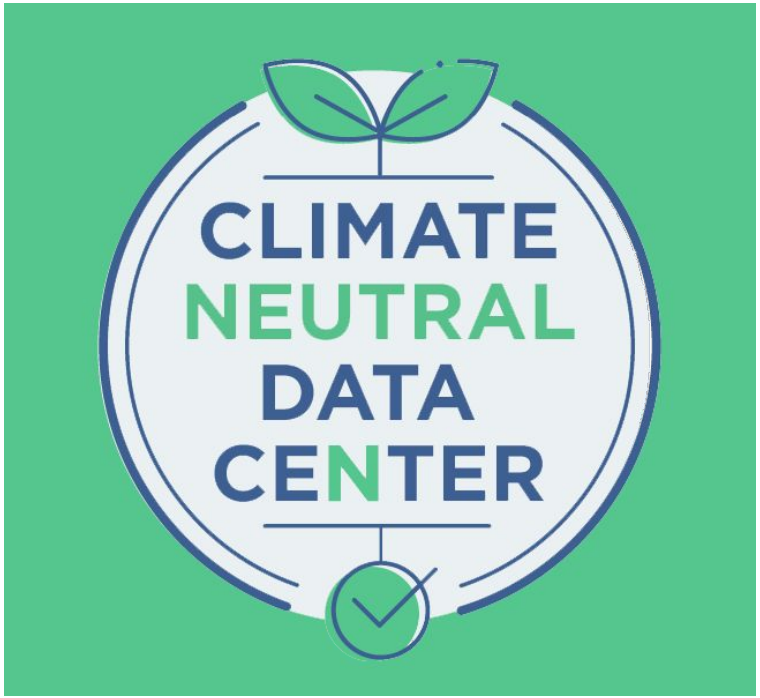
UCSDCSE
Computer Science and Engineering

E-lab

Datacenters account for increasing energy usage and carbon emissions



https://www.akcp.com/blog/the-real-amount-of-energy-a-data-center-use/



https://www.datacenterdynamics.com/en/analysis/european-cloud-providers-pledge-go-climate-neutral-2030/

Datacenters transitioning to a mix of energy sources

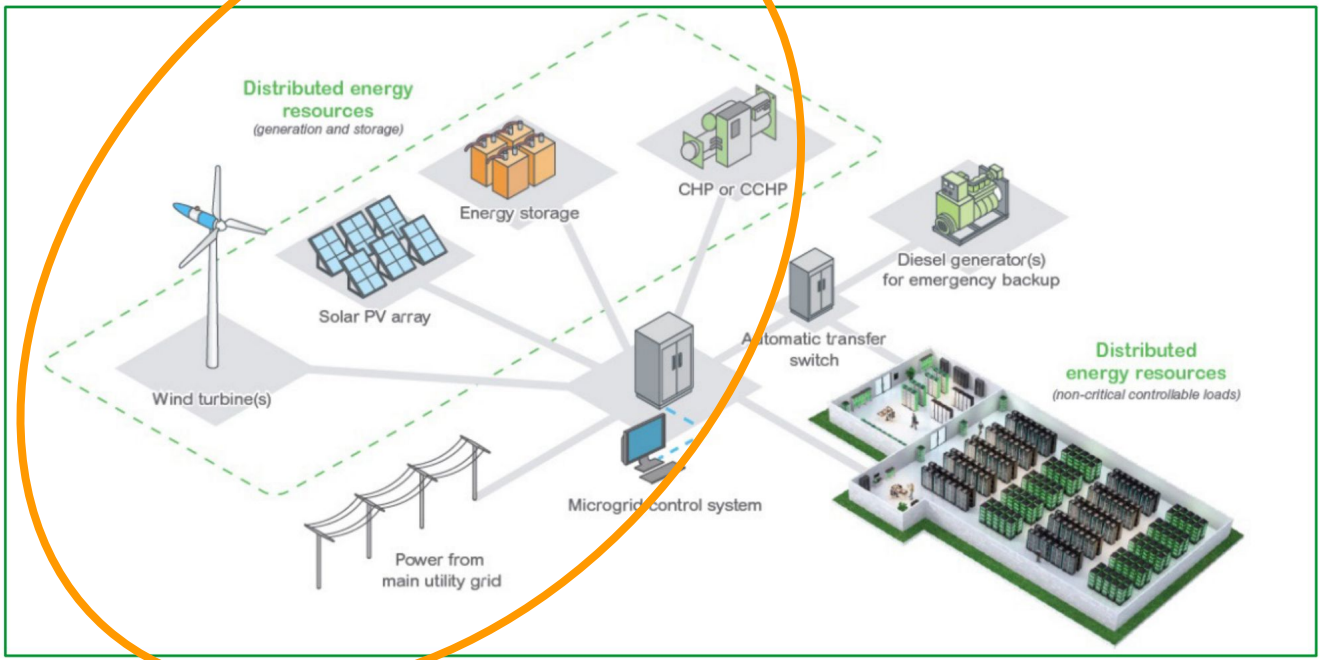
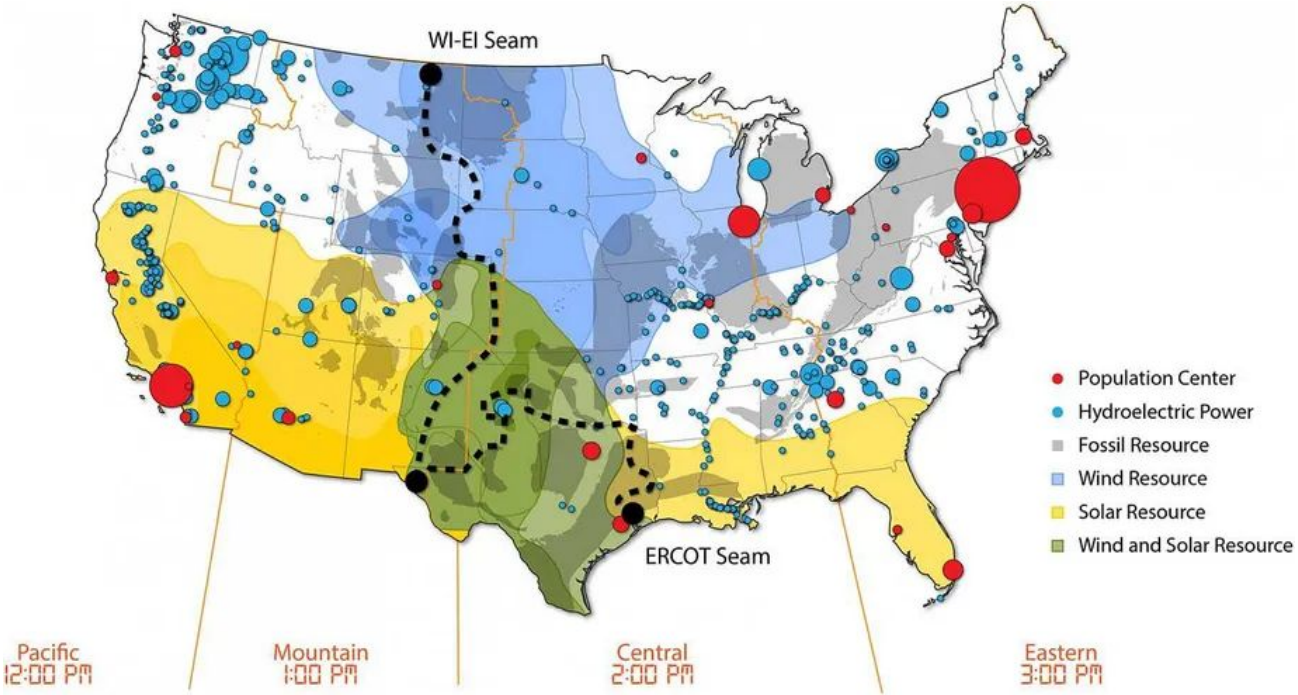


Figure 4, How Microgrids for Data Centers Increase Resilience, Optimize Costs, and Improve Sustainability, Schneider Electric

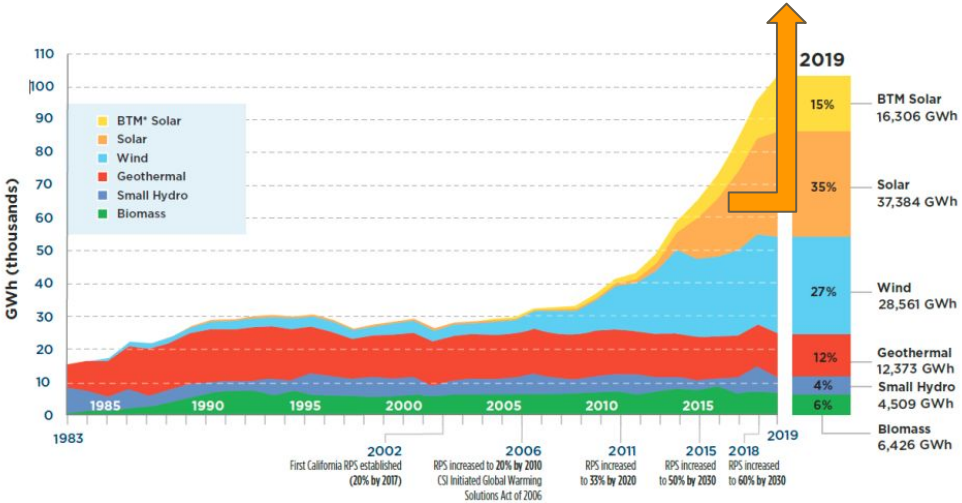
Low-carbon energy sources vary across time and space



Source: https://www.vox.com/energy-and-environment/2020/6/20/21293952/renewable-energy-power-national-grid-transmission-microgrids

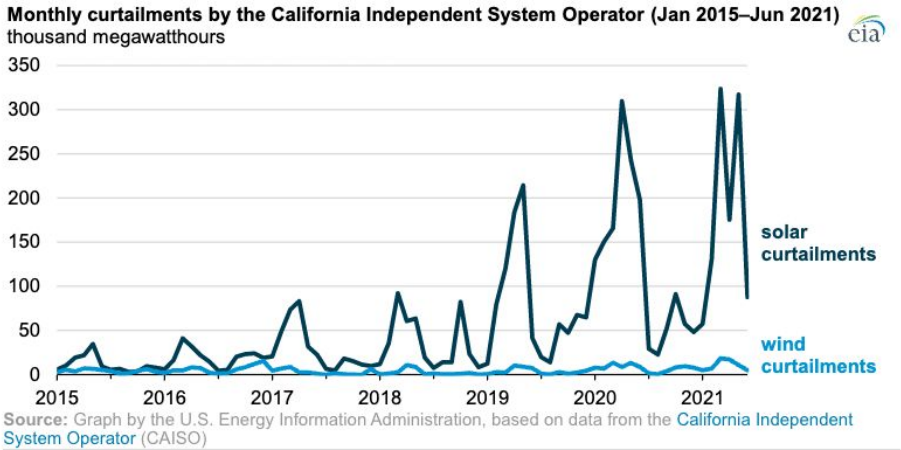
Increasing renewables but more challenges

California's increasing renewable deployment



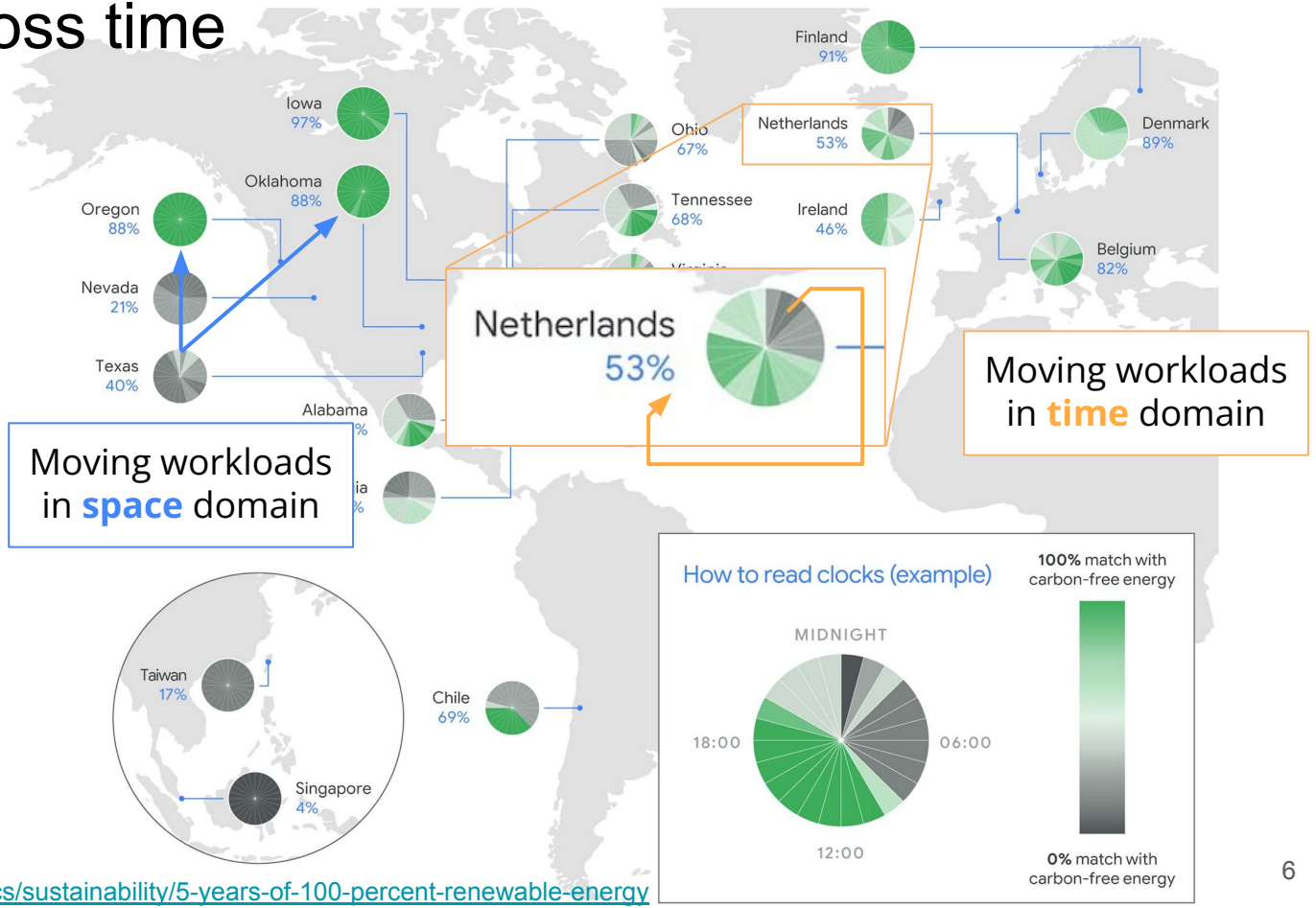
Source: energy.ca.gov: Renewable Tracking Progress, Dec 2019

More wasted solar/wind energy



Source: https://www.eia.gov/todayinenergy/detail.php?id=49276

Low-carbon energy sources vary across time and locations



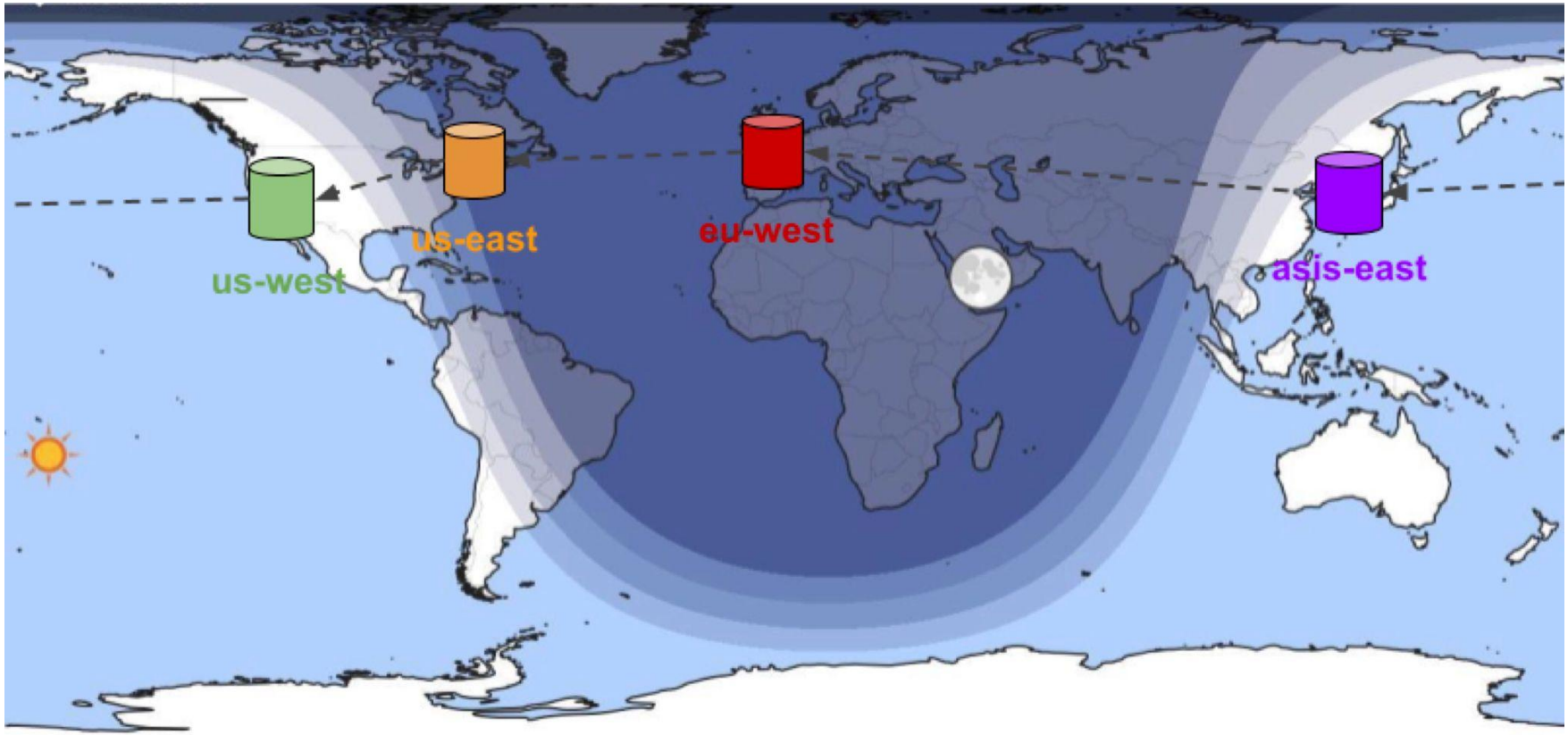
Source: https://cloud.google.com/blog/topics/sustainability/5-years-of-100-percent-renewable-energy

Focus not on reducing energy, but rather matching compute with low(er) carbon energy sources

Future computing: more varying renewables

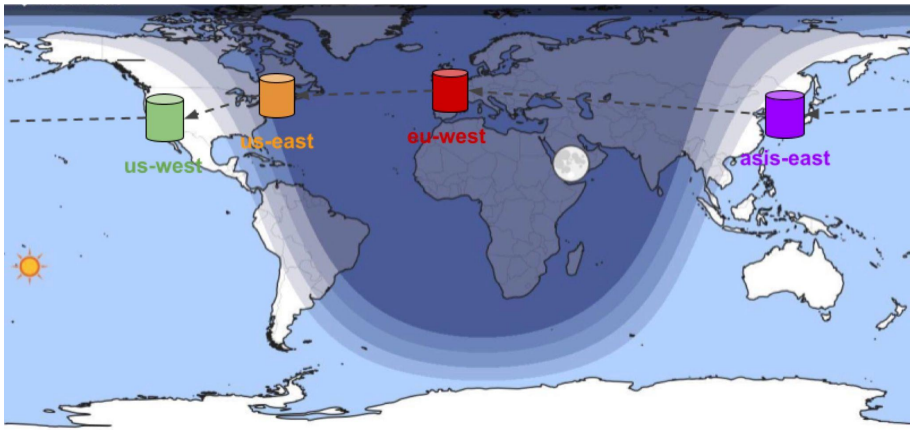
- The future grid will depend more heavily on these varying renewables
- Stronger need to shift from fixed power model to varying power model
  - Economic incentives to use excess renewables
  - Heavy investment in solar/wind energy
- Need to **make computing** as **flexible** as possible

What would a totally solar powered datacenter look like?



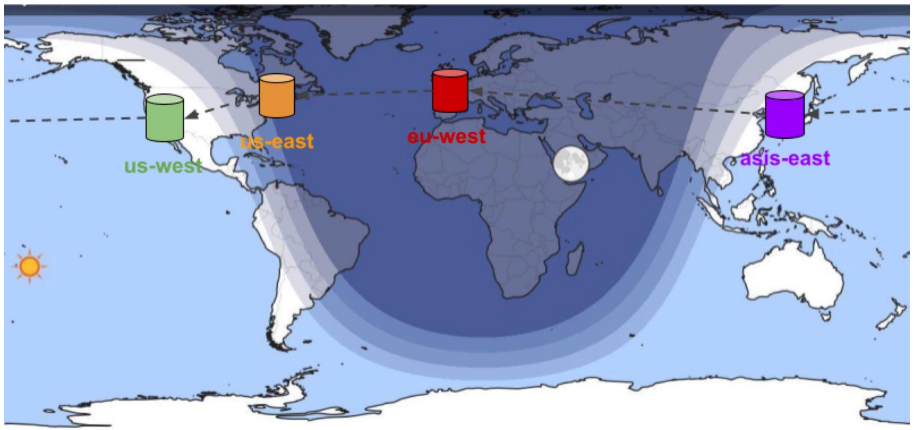
Challenges of solar-powered data centers

1. **How to schedule workloads across these data centers?**
    a. Time-shifting is not enough: high solar curtailment during daytime
    b. Space-shifting is possible: moving compute is cheaper than moving electricity
    c. Challenge: need to make compute more flexible



Challenges of solar-powered data centers

1. How to schedule workloads across these data centers?
2. **How to easily move workloads and what's the impact?**
    a. Latency-sensitive workloads: hard problem, e.g. QoS requirement
    b. Semi-flexible and batch workloads: slightly higher delay, but much lower carbon emissions
    c. How much overhead does moving a job incur?

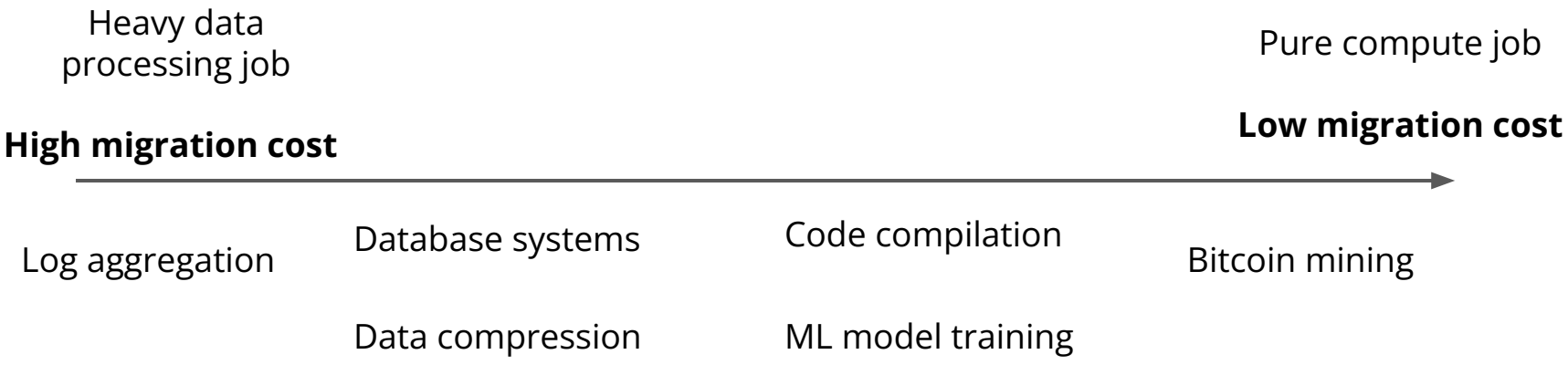


Metric of interest: Energy moved per unit of data moved

Jobs are computation over data

- Higher input/output data size means more migration overhead
- Higher compute usage offsets the migration overhead

***compute-to-data-size ratio* = Compute energy usage (CPU hour) / data size (GB)**

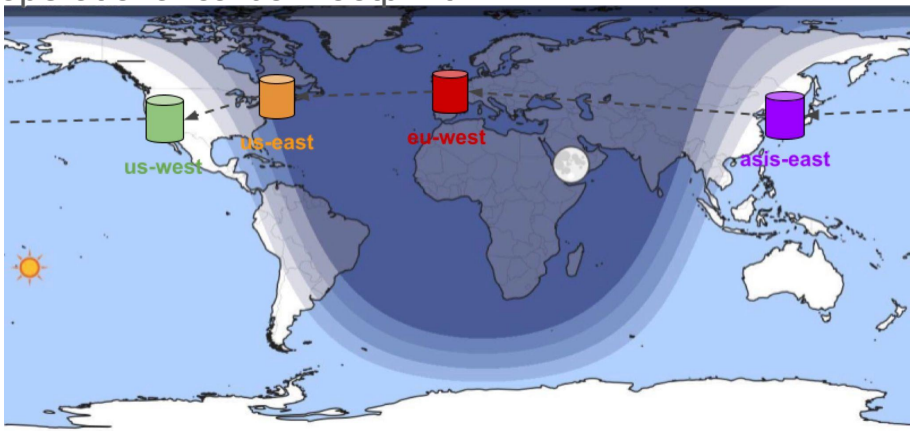


How to reduce overhead of moving a job?

- Prefer moving energy intensive jobs
  - Consume more clean energy and pay less migration cost
- Reduce the amount of data to move:
  - Replicated dataset for fault tolerance
  - Shared/common input dataset
  - Pause/restart for multi-day jobs
- Application tracing and lineage capture to map workloads to the datasets they use (and find applications which share datasets)
- Deploy additional WAN bandwidth

Challenges of solar-powered data centers

1. How to schedule workloads across these data centers?
2. What's the impact on each type of workload?
3. **How much additional capacity is needed?**
    a. Compute and network capacity
    b. Embodied vs operational carbon footprint



Capacity challenge and embodied carbon footprint

- Pure solar-powered data centers requires up to 3x servers, but
  - There are more stable renewables like wind/hydro/geothermal
  - We can keep old servers or overclock CPUs if we have excess solar power, to avoid additional embodied carbon footprint
- We need extra wide-area network bandwidth
  - But they are much cheaper than moving electricity via high-voltage lines
  - Existing studies have shown that high-speed transfer between data centers is possible, e.g. Skyplane *NSDI'23*

Summary

We can reduce operational carbon footprint by better utilizing renewables.

- Need for flexible computing: moving across time and space
- One metric to optimize: *Energy moved per unit of data moved*
- Careful profiling of workloads to reduce migration cost
- Balance between embodied and operational carbon footprint