Carbon Explorer: A Holistic Framework for Designing Carbon Aware Datacenters

Bilge Acun[1], Benjamin Lee[1,2], Fiodar Kazhamiaka[1,3], Kiwan Maeng[1], Manoj Chakkaravarthy[1], Udit Gupta[1,4], David Brooks[1,4], Carole-Jean Wu[1]

[1]Meta, [2]University of Pennsylvania, [3] Stanford University, [4]Harvard University

Abstract

Technology companies have been leading the way to a renewable energy transformation, by investing in renewable energy sources to reduce the carbon footprint of their datacenters. In addition to helping build new solar and wind farms, companies make power purchase agreements or purchase carbon offsets, rather than relying on renewable energy every hour of the day, every day of the week (24/7). Relying on renewable energy 24/7 is challenging due to the intermittent nature of wind and solar energy. Inherent variations in solar and wind energy production causes excess or lack of supply at different times. To cope with the fluctuations of renewable energy generation, multiple solutions must be applied. These include: capacity sizing with a mix of solar and wind power, energy storage options, and carbon aware workload scheduling. However, depending on the region and datacenter workload characteristics, the carbon-optimal solution varies. Existing work in this space does not give a holistic view of the trade-offs of each solution and often ignore the embodied carbon cost of the solutions.

In this work, we provide a framework, Carbon Explorer, to analyze the multi-dimensional solution space by taking into account operational and embodied footprint of the solutions to help make datacenters operate on renewable energy 24/7. The solutions we analyze include capacity sizing with a mix of solar and wind power, battery storage, and carbon aware workload scheduling, which entails shifting the workloads from times when there is lack of renewable supply to times with abundant supply. Carbon Explorer is open-sourced on Github.

1. Introduction

In 2021, the United Nations created a *24/7 Carbon-Free Energy (CFE) Compact* with a call for an ambitious goal: "Every kilowatt-hour of electricity consumption is met with carbon-free electricity sources, every hour of every day, everywhere" [65]. At the beginning of 2022, the U.S. Department of Defense and the U.S. General Services Administration requested information on strategies for "supplying 24/7 carbon pollution-free electricity for federal government" [66]. However, there is no clear pathway to achieve this 24/7 CFE vision for hyperscale datacenters (DC).

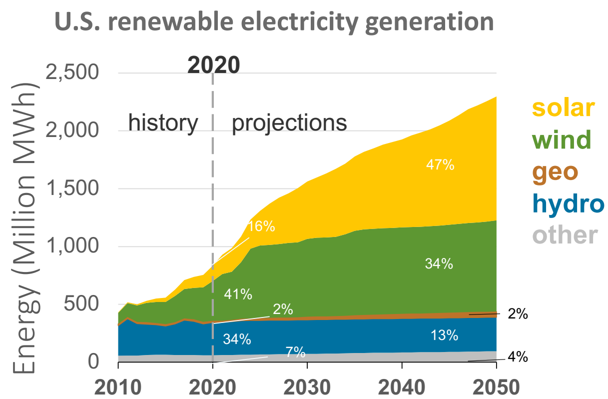

Figure 1: Historical and projection of growth of renewable energy in the US electricity grid [2010-2050]. 81% of the renewable energy will be comprised of solar and wind [67].

Datacenters world-wide are estimated to consume 205 TWh of electricity in 2018 [50], exceeding the annual consumption of countries such as Ireland and Denmark [58]. More generally, information and communication technology (ICT) is expected to account for 7% to 20% of the global electricity demand by 2030 [2, 27]. Technology companies mitigate computing's environmental footprint by investing in renewable energy generation to offset consumption by hyperscale datacenters [17, 28, 32]. Amazon, Meta, and Google have collectively invested in over 22 GW of renewable energy generation to meet their *Net Zero* commitments for datacenters and other operational activities. These investments in renewable energy for computing align with the broader trends in Figure 1. Renewable energy generation is projected to increase from 20% (in 2020) to 42% of the total by 2050 as the United States seeks to achieve its *Net Zero* goals [63]. Moreover, solar and wind comprise 47% and 34% of this renewable energy [67].

Despite these promising trends, computing with renewable energy twenty-four hours a day, seven days a week — 24/7 operational carbon-free computing — is challenging because renewable energy generation is highly intermittent. Figure 2 highlights the fluctuations in renewable

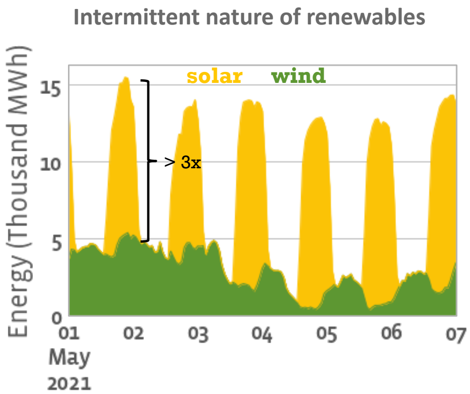

Figure 2: Hourly wind and solar energy generation in California grid during a week of time-frame.

energy generation in the California grid which has 33% share of renewables. The broad deployment of solar and wind farms will lead to increasingly severe hourly and seasonal fluctuations in energy generation. At times, the grid's supply of renewable energy may exceed demand, forcing inefficient curtailments that deactivate renewable energy generation in order to match supply with demand and reduce congestion on the power transmission network [5, 8, 9, 40]. At other times, solar and wind energy is scarce and datacenters risk consuming carbon-intensive energy from gas or coal from the grid despite their investments in renewable generation.

Under these conditions, datacenters can claim Net Zero operations on an annual basis, matching total renewable energy credits generated by its wind and solar investments against the energy consumed by its computation, yet continue to generate carbon emissions on an hourly basis due to fluctuations in wind and solar supply. Thus, more must be done to eliminate carbon emissions for datacenter computing in every hour of every day.

Various strategies have been explored to help datacenters cope with the intermittent nature of renewable energy and increase 24/7 CFE. First, demand response strategies can predict renewable energy supply and datacenter energy demand, and schedule computation to align supply and demand [4, 54, 71, 76]. Second, energy storage can reduce datacenter exposure to fluctuations in renewable energy supply [43, 46]. Third, further investment in wind and solar can ensure sufficient supply for more hours and days of the year. These strategies have significant implications for datacenter infrastructure. Datacenters may need additional servers that perform extra computation when carbon-free energy is abundant [77], batteries that provide capacity well beyond that of today's existing power supplies, as well as investments in energy generation that reflect the datacenter's location and the relative availability of wind and solar. All of this infrastructure incurs additional embodied carbon cost – an important but under-explored aspect of carbon-efficient datacenters.

In this paper, we present a framework for defining and exploring the design space for 24/7 carbon-free computing. The framework consumes a vast amount of data that details energy demands from hyperscale datacenters and energy supplies from renewable sources across locations in the United States. It models the effect of investments in renewable energy generation using real power grid data, in battery capacity using a physically accurate battery storage model, and in server capacity to support computation scheduling using production-datacenter traces. Finally, the framework performs a Pareto analysis to determine optimal carbon footprint designs while taking into account both the operational and the embodied carbon incurred when manufacturing the required infrastructures. In summary, the main contributions of the paper are:

- We propose and provide a new design space exploration framework — *Carbon Explorer* — that enables system architects to optimize environmentally-sustainable datacenters. Carbon Explorer models and optimizes datacenters to permit 24/7 carbon-free computing (Section 2).
- We characterize the hourly renewable supply and DC demand patterns for each region where *CompanyX* has datacenters at (Section 3). Out of the thirteen locations, we find that Nebraska (a majorly-wind region), Utah and Texas (wind and solar hybrid regions) datacenters are the best locations to minimize total carbon footprint because the lowest energy generation days throughout the year has higher renewable energy generated compared to other regions.
- We take a holistic approach in designing a datacenter to operate 24/7 on renewable energy by taking into account embodied carbon footprint, an important but under-explored area of the design space. This embodied carbon footprint comes from the additional server capacities, energy storage units, and renewable energy infrastructures (Section 4). When embodied carbon footprint is considered, carbon-optimal datacenter design is ***not*** always to reach 100% 24/7 renewable coverage.
- Carbon Explorer provides insights on how effective each strategy will be for each DC region:
  - With a *Renewable-Only* strategy, optimal renewable coverage ranges from 46% to 99%. With a *Renewable+Battery* solution, deploying sufficient batteries to reach 100% coverage is the carbon-optimal solution for nine DC regions. For the remaining four DCs, 99% 24/7 carbon-free coverage is the optimal (Section 5).
  - *Renewables+Scheduling* increases 24/7 coverage by 1%-21% for different regions. Carbon Explorer shows that deploying 6% to 76% additional server capacity to allow for carbon aware scheduling reduces the overall carbon footprint of the datacenters given

sufficient workload flexibility (Section 5).

Climate change is an existential crisis. We hope Carbon Explorer will enable future works to deploy carbon-optimal technologies and achieve environmentally-sustainable computing in the years to come.

2. Carbon Explorer

Figure 3 illustrates *Carbon Explorer*, a design space exploration framework that takes a holistic approach to achieve 24/7 carbon-free computing. Carbon Explorer considers two important inputs: time-series data that details the power demand of large-scale datacenters and the intermittent nature of renewable energy generation at specific geographic locations (Figure 3-left). Next, Carbon Explorer characterizes a solution space that spans the following dimensions (Figure 3-center):

- Investments in varied types of renewable energy,
- Investments in varied amounts of energy storage,
- Scheduling that shifts varied amounts of computation.

Finally, Carbon Explorer models the datacenter design space and minimizes the carbon footprint, accounting for both operation **and** embodied carbon (Figure 3-right).

Carbon Explorer's inputs include two hourly time series. The first details power consumed by each of *CompanyX*'s datacenters in various locations across the United States. The second details energy generation for balancing authorities (BAs) in each datacenter location. Table 1 summarizes these locations and *CompanyX*'s renewable energy investments. Section 3 characterizes datacenter energy demand and renewable energy supply, which has implications for a production datacenter's carbon footprint.

We evaluate three distinct solutions for 24/7 carbon-free datacenters. First, datacenters could offset their energy consumption with renewable energy generation. Operators invest in wind and solar farms on the power grids that supply their datacenters. Moreover, they implement power purchase agreements, which issue credits for renewable energy generated from those investments and offset datacenter energy consumed. This state-of-the-art solution has been central to hyperscale datacenters' pursuit of Net Zero goals.

Second, datacenters could install energy storage and batteries to handle the intermittent availability of renewable energy. Although today's datacenters do not yet deploy batteries to manage their operational carbon footprint, they do deploy batteries to ensure system resilience and shave power peaks [33, 48]. As lithium-ion batteries mature, they become cost-effective for deployment at scale. On-site energy storage enables a new strategy for 24/7 carbon-free datacenters.

Finally, datacenters could schedule computation in response to renewable energy supply. Such demand response likely requires investment in additional servers. A datacenter that defers tasks when renewable energy is scarce must compute for those tasks when renewable energy is abundant, generating demand for servers above and beyond typical loads. In effect, bursts of renewable energy generate bursts of computation and demand for servers.

The three solutions lead to trade-offs between operational and embodied carbon footprints. Renewable energy permits carbon-free operation but is constrained by energy availability and consistency, which varies with geography. Large batteries store carbon-free energy but incur carbon overheads from manufacturing. Additional servers permit demand response and scheduling but also incur carbon overheads from manufacturing [24].

Carbon Explorer defines a comprehensive design space for 24/7 carbon-free datacenters. Section 4 navigates trade-offs in the solution space with a quantitative approach. And Section 5 illustrates the solution space for various geographic locations, highlighting the impact of site selection for future datacenters.

3. Operational Grid Inputs: Demand and Supply Characteristics

Carbon-aware datacenter design requires understanding datacenter energy demand and renewable energy supply, at fine granularity and in every region. Table 1 lists the locations of *CompanyX*'s datacenters and renewable energy investment, each identified by the balancing authority for the electric grid [10]. Total investment in renewables is nearly six Gigawatts.

Table 1: *CompanyX*'s Datacenter Locations in the U.S. and Regional Renewable Investments [10]

Location	Balancing Authority	Renewable Investment [MW]		
		Solar	Wind	Total
1. Sarpy County, Nebraska (NE)	SWPP	0	515	515
2. Prineville, Oregon (OR)	BPAT	100	0	100
3. Eagle Mountain, Utah (UT)	PACE	694	239	933
4. Los Lunas, New Mexico (NM)	PNM	420	215	635
5. Fort Worth, Texas (TX)	ERCO	300	404	704
6. DeKalb, Illinois (IL)	PJM			
7. Henrico, Virginia (VA)	PJM	840	309	1149
8. New Albany, Ohio (OH)	PJM			
9. Forest City, North Carolina (NC)	DUK	410	0	410
10. Altoona, Iowa (IA)	MISO	0	141	141
11. Newton County, Georgia (GA)	SOCO	425	0	425
12. Gallatin, Tennessee (TN)	TVA	742	0	742
13. Huntsville, Alabama (AL)	TVA			
	Total	1823	3931	5754

Our energy supply analysis draws data from the U.S. Energy Information Administration (EIA) Hourly Grid

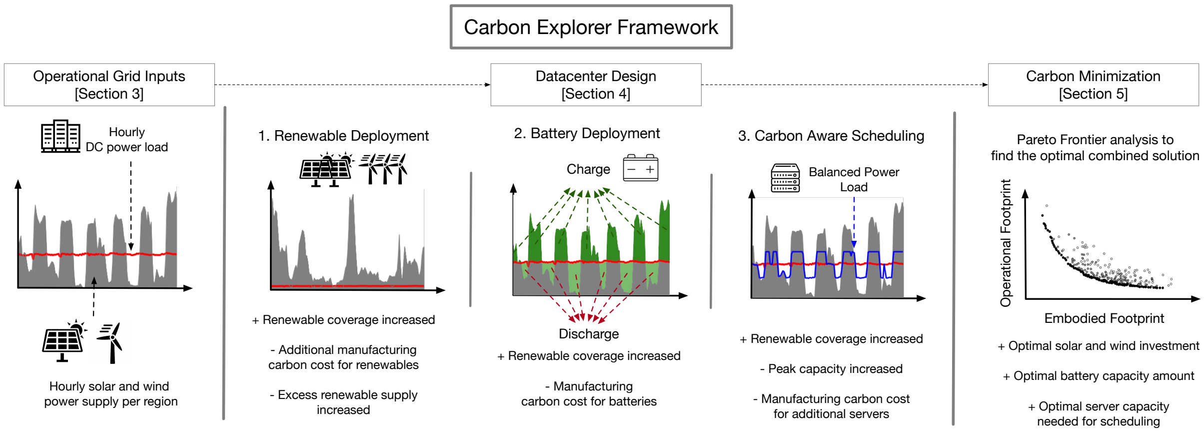

Figure 3: Design Overview for Carbon Explorer. Carbon Explorer considers important characteristics, such as time-series power demand of large-scale datacenters and renewable energy availability on the power grids, as inputs. Carbon Explorer characterizes the design space across renewable energy investments, energy storage, and computation shifting. Carbon Explorer provides *quantitative measures for strategies to achieve carbon-minimum settings*.

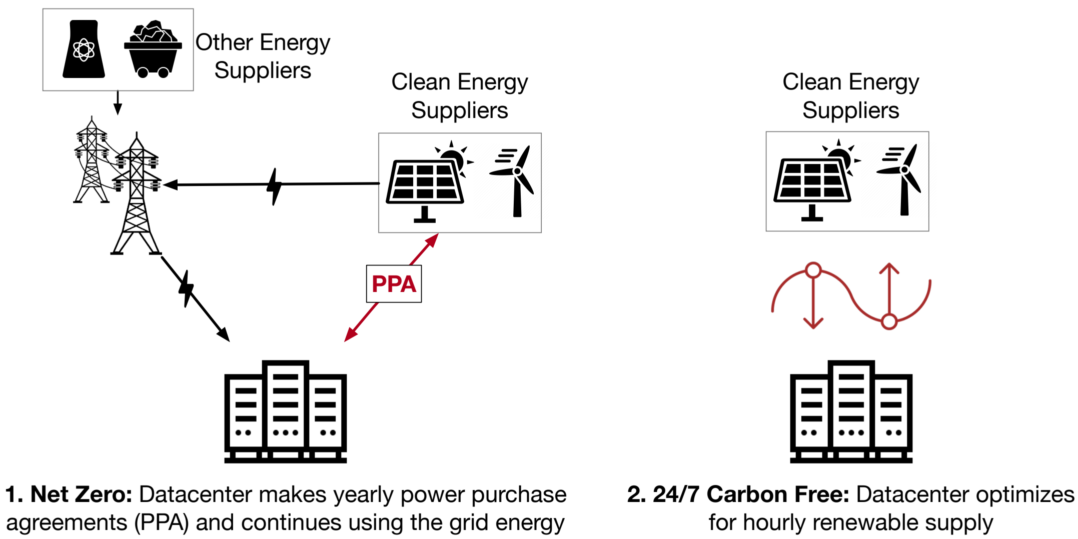

Figure 4: Illustration of datacenter power supply scenarios: (1) Net Zero Using Offsets, (2) 24/7 Carbon Free.

Monitor, which provides operating data for power grids in the lower 48 states [68]. Launched in 2019, the monitor provides hourly generation statistics by collecting data from balancing authorities (BAs). Each BA operates a grid and balances electricity flows, controlling electricity generation and transmission within its own region and between neighboring authorities.

3.1. Characterizing Datacenter Power Demand

CompanyX has built hyperscale datacenters across the globe with different capacities. These datacenters exhibit diurnal load patterns due to variations in user activity and exhibit peaks due to special events and holidays. Figure 5 shows diurnal usage for *CompanyX* and Google datacenters and illustrates how power usage correlates with processor utilization. *CompanyX*'s CPU utilization and power is averaged over three months. For *CompanyX*, CPU utilization swings by about 20% for an average datacenter and can swing by even more for an individual datacenter. For Google, the difference between the maximum and the minimum CPU utilization is 15%, on average [64].

However, diurnal patterns from interactive computation do not translate directly into power patterns. At datacenter scale, the difference between maximum and minimum energy demand is around 4%, on average, which is relatively insignificant compared to the swings in renewable energy supply. Thus, in today's datacenters, power variations will arise primarily from supply but not demand. Yet shifting computation to modulate datacenter power is possible because workloads exhibit different flexibility levels and come with distinct service level objectives (SLOs). The highest priority, user-facing services require real-time response. Latency-tolerant workloads, such as

Table 2: Operational Carbon Intensity of Energy Sources

Type	gCO_2eq/kWh	Type	gCO_2eq/kWh
Wind	11	Natural Gas	490
Solar	41	Coal	820
Water	24	Nuclear	12
Oil	650	Other (Biofuels etc.)	230

batch and AI training jobs [61, 73], target specific SLO categories that include 4-, 8- and 24-hour completion times. Google has reported that flexible jobs with 24-hour completion SLOs make up about 40% of the Borg scheduler's jobs [64]. This flexibility permits carbon-aware workload scheduling.

3.2. Characterizing Datacenter Power Supply

Figure 6 presents two scenarios that describe how datacenters could consume energy from rapidly evolving power grids. In the first scenario, datacenter operators collaborate with utility providers to invest in renewable energy on the grids that power the datacenters by purchasing energy with sophisticated accounting frameworks that track renewable energy credits. This represent the state-of-the-art in reducing a datacenter's operational carbon footprint for Net Zero commitments [17, 28, 32]. In the second scenario, datacenter operates on renewable energy 24/7 by optimizing hourly supply and demand.

Net Zero. Datacenter operators invest in renewable generation, such as wind and solar, and implement power purchase agreements (PPAs) to reduce datacenter exposure to the grid's carbon intensity. PPAs link renewable energy credits (RECs) with a specific source of energy and issue, e.g., one certificate for every MWh generated [16, 20, 29].

With RECs, the energy consumed is much greener than the energy offered by the grid. Table 2 details the carbon intensity of different electricity sources in the grid. The grid's energy mix is determined by the utility provider's dispatch stack and portfolio of generating assets [7]. But the datacenter's energy mix is determined by its pre-negotiated PPAs, which deliver carbon-free energy. Given datacenter operators' investments in renewable energy, most energy consumption may be matched and therefore is carbon-free but, during the remaining times, energy consumption is as carbon-intensive as the grid supply [21]. Figure 6 illustrates an example of how PPAs enable Net Zero computing. Wind and solar energy generation varies across days even as datacenter energy consumption is relatively constant. Renewable energy credits are issued as wind and solar energy is generated. At the end of the month (and end of the year), the total amount of energy generated and credits issued is equal or greater than the total amount of energy consumed. Thus, datacenters achieve a Net Zero carbon footprint on a monthly or annual basis. Although on an hourly basis, the carbon intensity of the energy used can be as much as the grid's carbon intensity during the times when there is not enough renewable supply (i.e. white areas under the red line in Figure 6).

24/7 Carbon Free. In addition to installing renewable energy, 24/7 carbon-free datacenters must address variable, intermittent generation. Figure 8 highlights variability across geography with rows corresponding to three representative regions with distinct renewable energy profiles: (a) Oregon BPAT with wind; (b) North Carolina DUK with solar; (c) Utah PACE with a mix of wind and solar. More broadly, of the ten balancing authorities in Table 1, three offer primarily wind energy (BPAT, MISO, SWPP), three offer primarily solar energy (DUK, SOCO, TVA), and four offer a mix (ERCO, PACE, PJM, PNM).

Figure 8 also highlights variability across time with columns corresponding to summary statistics calculated over the year: (a) Yearly Average; (b) Highest 10 Days; (c) Lowest 10 Days [68]. On average, wind and solar installations provide significant supply, but averages obscure high variance across time. For BPAT, the best ten days of the year offer approximately 2.5× more renewable energy than the average whereas the worst offer very little. Histograms in Figure 8-(d) quantify this variance and illustrate uncertainty in wind and solar supply.

As solar and wind farms proliferate, peaks and valleys in energy supply will become increasingly extreme. Utility providers will find it increasingly difficult to match its supply to consumer's demand. For example, California's renewable sources can generate much more electricity than needed in the middle of the day [5]. And curtailments are needed to manage excess supply and reduce renewable energy generation [8, 9, 40]. Figure 7 indicates that, since 2015 the curtailed gap between supply and demand has grown steadily as wind and solar capacity has increased. In 2021, curtailments reached 6% of the total generated renewable energy in the California grid, which has deployed a significantly more renewable electricity compared to the U.S. average (33% vs 20% in 2020 [6, 62]).

It is becoming increasingly complicated to fully consume peak renewable energy generation due to the high variance. When supply exceeds demand, only generators with the lowest prices can supply energy to the grid. Prices can be zero or even negative because inputs to wind/solar farms are free and generators often receive government subsidies [14, 78]. As a result, grids may offer lower time-of-use energy prices and incentivize datacenters to defer computation to periods of abundant renewable energy.

Challenges in variable supply and curtailments require energy storage and demand response scheduling during periods of scarcity. Energy storage mitigates supply variations by providing carbon-free energy when solar and wind cannot [52, 53]. Demand response modulates dat-

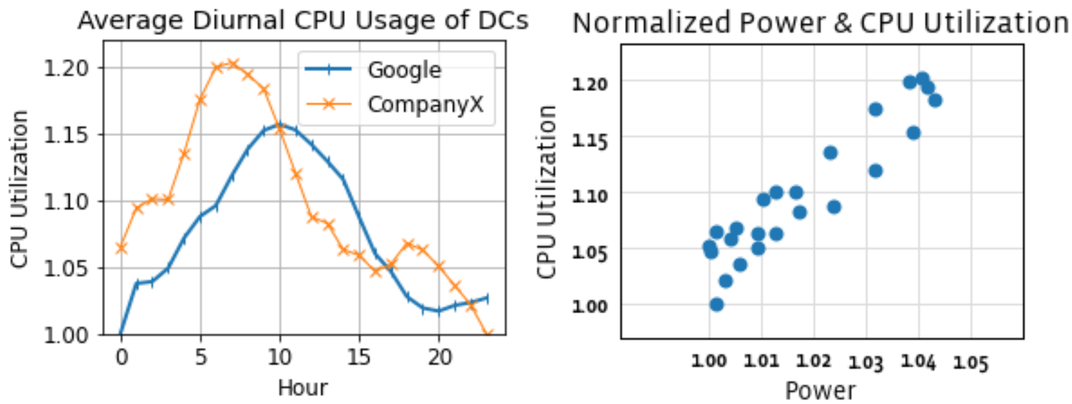

Figure 5: [left] Hourly DC CPU fluctuations of *CompanyX* and Google DCs. [right] Hourly CPU Utilization and Power correlation of *CompanyX* DCs.

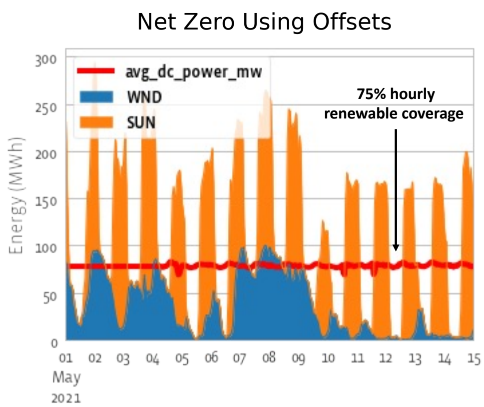

Figure 6: Datacenter power demand and corresponding renewable investments to achieve Net Zero operational carbon free.

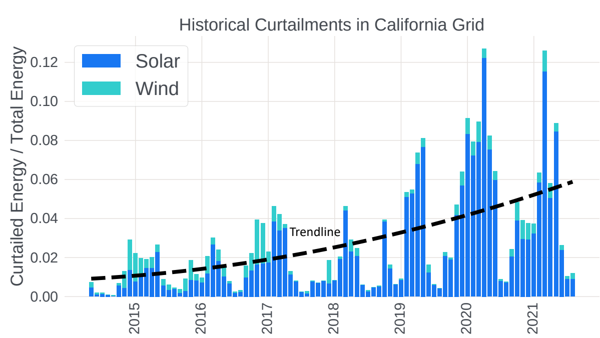

Figure 7: Wind and solar curtailments have been increasing with the renewables on the California grid [5].

acenter demand for energy based on signals about renewable energy supply. Signals could come in the form of utility surcharges or credits when the datacenter consumes or reduces its energy demand during various times of day [47]. Signals can also come from utility providers' generation statistics that describe the mix of green and brown energy across time and geographic locations. The most informative signals would communicate hourly variations in energy demand and supply.

In summary, the Net Zero scenario describes how renewable energy investments significantly reduce the carbon intensity of datacenter operations. And the 24/7 scenario describes how additional investments in energy storage and demand response schedulers could further reduce carbon intensity. Figure 9 compares the carbon intensity of these scenarios with that of the grid's energy mix. Next, in Section 4, we use Carbon Explorer to show how coordinated strategies for deploying renewable energy generation, energy storage, and demand response scheduling could lead datacenters to carbon-free operations on an hourly basis.

4. Datacenter Design: Strategies for Carbon Free Computing

A datacenter must implement a portfolio of complementary solutions to achieve its goal of using 24/7 carbon-free energy efficiently and robustly. Carbon Explorer considers renewable energy investments (Section 4.1), energy storage installations (Section 4.2), and carbon-aware scheduling (Section 4.3). In this section, we model and analyze these solutions and associated trade-offs in operational and embodied carbon footprints.

4.1. Renewable Energy

Carbon Explorer determines the solar and wind investments required for datacenters in different geographic regions to increase and achieve 100% hourly renewable coverage. We define *renewable coverage* as the percentage of hours in the year where datacenter power (P_{DC}) is covered by renewable power (P_{Ren}):

$$\left\{ \sum_{\text{hour}} \{P_{DC} - P_{Ren}\} / \sum_{\text{hour}} P_{DC} \right\} \times 100 \quad \forall \text{hour} \in \text{DateRange}$$

Carbon Explorer projects hourly wind and solar energy supply by scaling EIA grid data in proportion to the desired renewable investment level. It takes the maximum

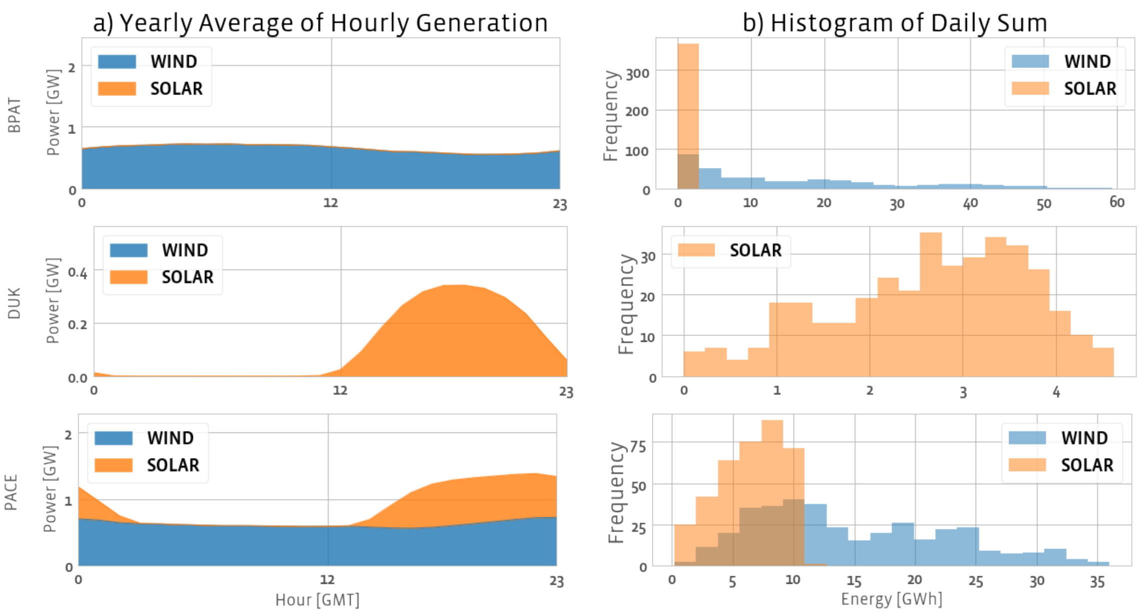

Figure 8: Figure shows hourly wind and solar generation of an average day in year 2020 (left) and histogram of daily sum to highlight day-to-day fluctuations of wind and solar generation in BPAT (in OR), DUK (in NC) and PACE (in UT) balancing authorities which are composed of majorly wind, solar-only and a mix of wind and solar energy correspondingly. The data is calculated over the entire year of 2020.

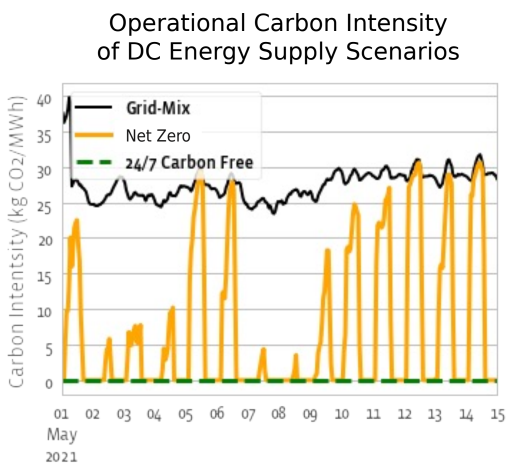

Figure 9: Comparison of hourly operational carbon intensity of different DC energy supply scenarios.

generated solar and wind power throughout the year as the maximum capacity of the local grid. Then, the hourly generation data is linearly scaled to the desired renewable investment capacity. Finally, hourly renewable supply data is matched against hourly datacenter demand for every region to calculate renewable coverage. Figure 10 shows renewable coverage (z-axis) with different wind (x-axis) and solar (y-axis) investments from two regions served primarily by wind and solar, respectively.

Figure 10 reports *CompanyX*'s existing renewable investments with black lines. While these investments help *CompanyX* achieve Net-Zero goals on a monthly or annual basis, coverage on an hourly basis is only 46% and 51% in the two regions. Each region tends to favor a particular type of renewable energy generation on its local grids and *CompanyX*'s investments generally align with those profiles. One exception is Oregon, where *CompanyX*'s investments emphasize solar despite the local grid's emphasis on wind.

For regions served primarily by wind energy, like Oregon, high day-to-day fluctuations increase the investment in wind generation needed to satisfy minimum energy needs. For regions that rely entirely on solar for renewable energy, it is impossible to increase 24/7 coverage much beyond 50% because solar energy is available only during the day. For regions that deploy a mix of solar and wind generated renewables, the tail is shorter and diminishing marginal returns in 24/7 coverage are less severe since wind and solar availability can complement each other.

There is a long tail to reach 100% renewable coverage. As coverage increases, curves flatten and indicate diminishing marginal returns from further investment in renewable generation. Figure 11 highlights the full length of the tail for Oregon's datacenter. It takes more than 5× more investments in renewable energy generation to go from 95% to 99.9% than to go from 0% to 95% coverage. Due to space limitations, we are unable to show profiles of every datacenter location. But this representative analysis shows that other solutions, such as energy storage and carbon-aware scheduling, are essential to complement renewable energy generation.

Note that accurate hourly energy supply data is cru-

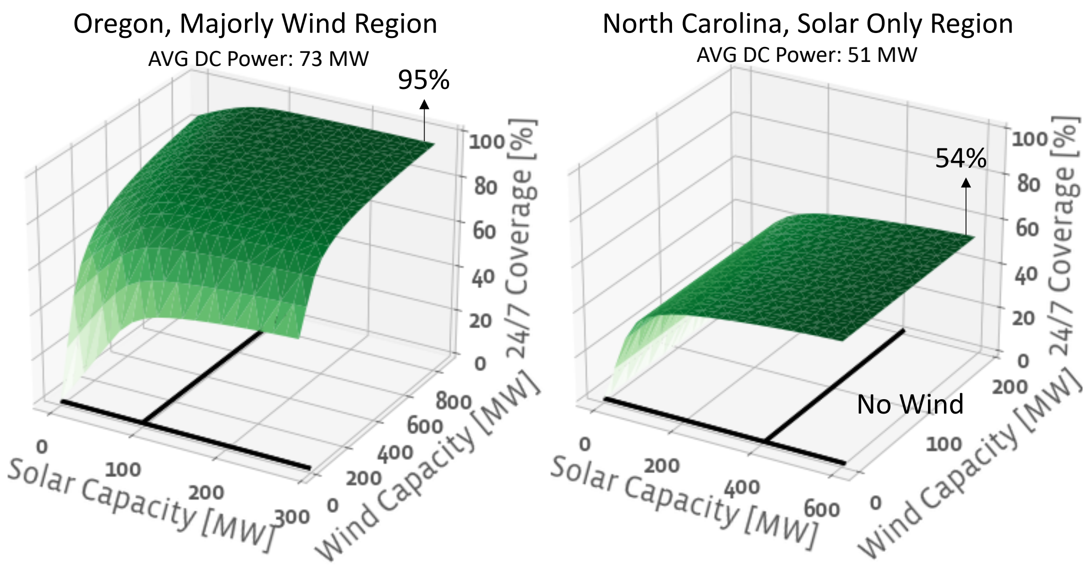

Figure 10: 24/7 coverage with varying amount of wind and solar investments. Black lines show *CompanyX*'s renewable investment amount in the corresponding region.

cial when making design decisions. Figure 11 shows that assuming wind and solar output to be same as average output every day leads to overly optimistic design conclusions. Under this assumption, achieving 100% coverage would require an order of magnitude less renewable investments. Thus, Carbon Explorer requires fine-grained time series supply and demand data when determining investments in renewable energy generation.

4.2. Battery Storage

Improvements in energy storage over the last decade have led to lithium-ion batteries (LIB) that offer high capacity and energy density [12]. As the technology has matured, LIB has become a common, cost-effective storage medium for renewable energy [74]. For these reasons,

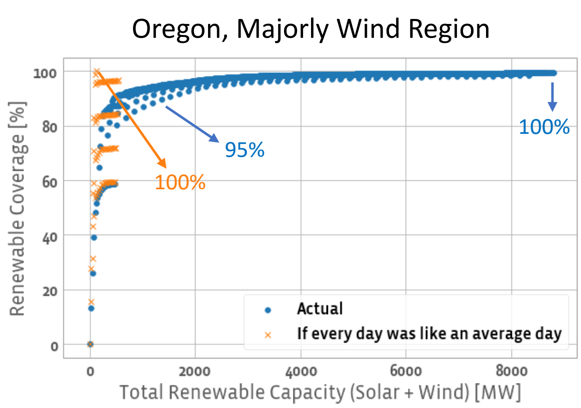

Figure 11: 24/7 coverage with different renewable investments highlighting the long tail. Each point represents a different solar + wind capacity combination.

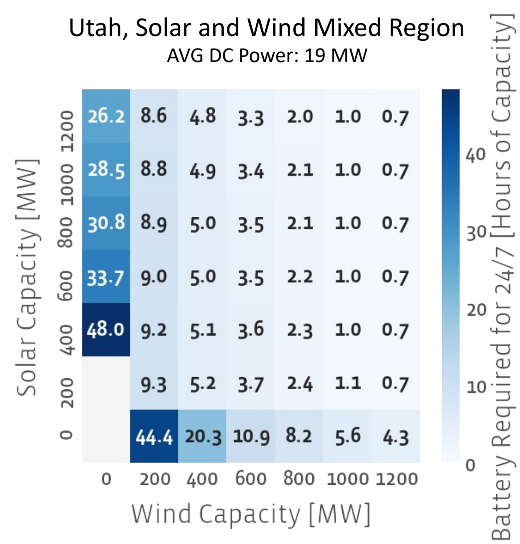

Figure 12: How much battery needs to be deployed for 24/7 renewable energy?

batteries play an important role in 24/7 carbon-free computing. This section describes how Carbon Explorer evaluates the impact of batteries that are charged by renewable energy and discharged by datacenter servers.

Datacenters already deploy batteries to prevent the interruption of services during maintenance or power failures. Batteries distributed throughout racks and clusters permit continuous operation when utility power fails and the datacenter must switch to diesel generator backups [49].

We envision batteries deployed on-site with the datacenter to reduce its carbon footprint. Batteries will be charged when there is excess renewable supply (i.e. when the amount of energy produced by the renewable deployment is larger than datacenter's demand). Batteries will be discharged to power the datacenter when there is a lack of renewable supply (i.e. when the amount of energy produced by the renewable deployment is smaller than datacenter's demand).

The battery model used in Carbon Explorer is the C/L/C model [31]; it explicitly models several characteristics of lithium-ion batteries, including energy content limits, efficiency loss, and limits on the applied power with respect to the energy content. Model parameters are tuned to represent a battery composed of Lithium Iron Phosphate cells [1] – a cell type found often in large stationary storage applications.

Figure 12 shows the amount of energy storage capacity required to reach 24/7 renewable energy coverage at different solar and wind capacities for Utah datacenter. Capacity is reported in terms of computational hours (e.g., 2 hours for a 20MW datacenter corresponds to 40MWh of battery capacity). Regions with mixed solar and wind generation exhibit less variable, day-to-day fluctuations and can achieve 24/7 carbon-free compute with less battery capacity. By adding around five hours of battery capacity to its existing renewable investments in the region, *CompanyX*'s Utah datacenter can reach 24/7 carbon-free operational energy. A battery of this size would be comparable to a utility-scale battery; the largest utility-scale energy storage project so far can offer 300 MW of power and 1,200-MWh of capacity [13].

In contrast, battery capacity requirements for 24/7 are greatest for regions that rely majorly on wind. Oregon suffers from extremely high day-to-day fluctuations and there are days with almost no wind power. Requirements are also high for regions that rely entirely on solar. For example, North Carolina datacenter requires 14 hours of battery-based compute.

4.3. Carbon Aware Scheduling

Carbon-aware scheduling (CAS) exploits delay tolerant workloads to achieve 24/7 carbon-free computing, shifting workloads from times when the carbon intensity of electricity sources is high to times when it is low. Hyperscale datacenter workloads are commonly organized into tiers based on their Service Level Agreements (SLAs). Higher tier jobs are latency sensitive and require high availability.

On the other hand, lower tiers can tolerate delays. Examples temporally flexible workloads include AI model training, data processing pipelines, and offline video processing. Google traces indicate a significant fraction of jobs submitted to the Borg scheduler are in the free and best-effort-batch tiers with weak SLAs [64]. Jobs at *CompanyX* exhibit similar characteristics — 60% of batch jobs must either land with daily consistency or their landing time is not important.

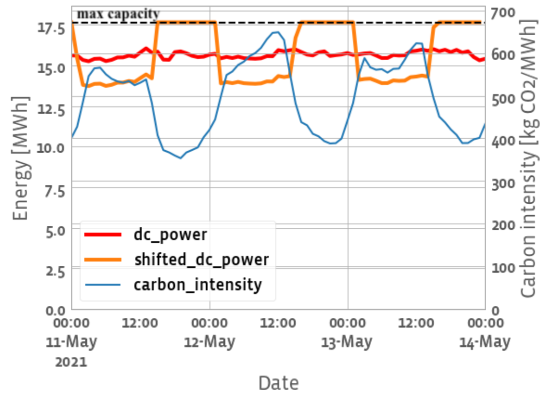

Figure 13: Carbon aware scheduling illustration for the Utah DC.

Carbon Explorer estimates the potential benefits of carbon aware workload scheduling using a greedy algorithm. The algorithm takes two customizable input constraints: datacenter capacity and flexible workload ratio for each hour of the day. Given these two constraints, flexible workloads are moved from times of highest carbon intensity to times of lowest intensity until all flexible workloads have been moved or all datacenter servers have been used for the given hour.

Input Constraint 1: $P_{DC_{MAX}} = Maximum\ DC\ Capacity$
Input Constraint 2:
$FWR = Flexible\ Workload\ Ratio\ (\%)$
Goal: For each day, minimize:

$$\sum_{h \in hour} \{P_{DC}(h) - P_{Ren}(h)\}$$

where
$P_{DC}(h) < P_{DC_{MAX}}$ and
$P_{DC}(h) \times FWR$ is allowed to shift

Figure 13 illustrates an example of a carbon aware scheduling over three days. The blue line shows how the grid's carbon intensity varies depending on the hour of the day. The red and orange lines shows datacenter power draw when carbon aware scheduling is and is not applied. In this example, the maximum allowed power capacity of the DC is assumed to be 17.6 MW and 10% of the workloads running every hour are flexible to finish within a day.

Additional Servers. Shifting computation across time may require additional server capacity for sustained increases in computation when carbon-free/low-carbon energy is abundant. The need for surplus capacity reveals an interesting trade-off between operational and embodied carbon. From an operations perspective, increasing the number of provisioned servers mitigates the data center's carbon footprint by permitting demand response and

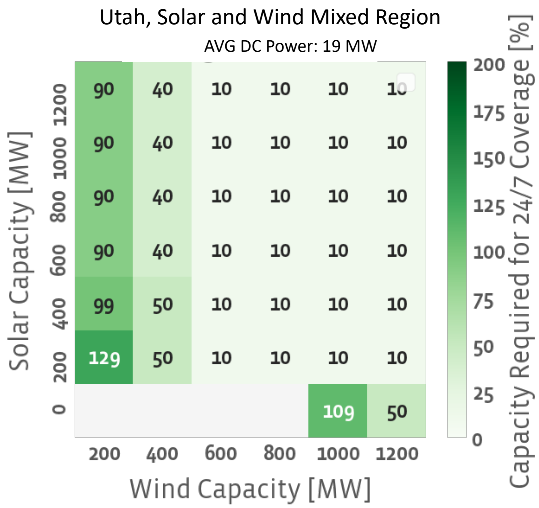

Figure 14: Scenario 3: Carbon aware scheduling for 24/7.

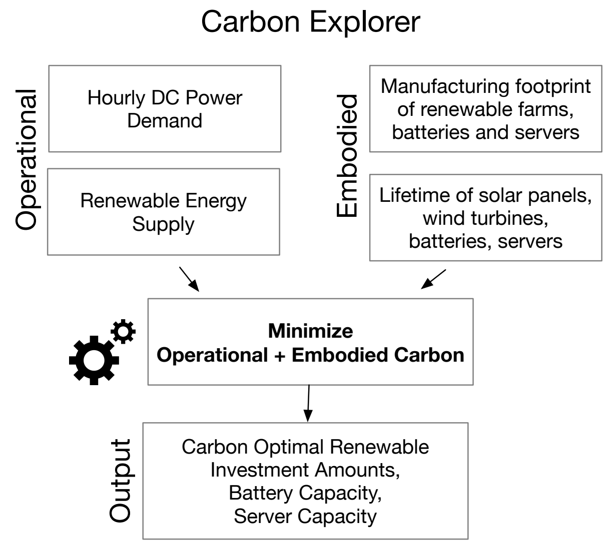

Figure 15: Carbon Explorer

reducing the carbon intensity of its energy. However, from the embodied perspective, over-provisioned servers increase embodied carbon emissions from hardware manufacturing [24]. Therefore, there is a fine balance between operational and capital expenditures.

Energy-proportional design is essential when over-provisioning datacenter servers [3]. Idle servers should draw little power, especially since hourly scheduling decisions provide ample time for servers to switch between power models. Indeed, server power can be accurately modeled as a linear function of utilization with the y-intercept denoting a server's idle power. Figure 5 illustrates energy-proportionality and correlation between *CompanyX*'s datacenter power and its CPU utilization.

Figure 14 shows how much server capacity is required to achieve 24/7 carbon-free computation. Additional capacity is measured as a percentage of the datacenter's existing capacity. In this example, all workloads are assumed to be flexible to shift. Analysis shows that the additional capacity required to reach to 24/7 varies between 19% to over 100% (i.e. doubling the number of servers). Note that, as an alternative to deploying more servers, datacenters might Turbo Boost their current servers to increase compute throughput without increasing capital costs and embodied carbon.

5. Carbon Minimization: Holistic Design Exploration

Reaching 24/7 carbon-free computing comes with non-negligible embodied carbon costs. Thus, Carbon Explorer must consider *both* operational and embodied carbon when minimizing the overall carbon footprint. Figure 15 presents the process of identifying an optimal datacenter design point from the carbon footprint's perspective.

First, Carbon Explorer requires inputs for its models of operational and embodied carbon. Operational inputs include hourly datacenter power demand and renewable power supply. Embodied inputs account for the carbon emissions from manufacturing and the expected lifetimes of solar and wind farms, lithium-ion batteries, and datacenter servers.

Carbon Explorer exhaustively searches the design space to minimize the sum of operational and embodied carbon. The design space includes the three solutions — renewable, battery, server investments — described in detailed in Section 4. Datacenter operators specify the bounds of the design space. Finally, Carbon Explorer outputs the carbon-optimal investments in renewable energy generation, battery capacity, and server capacity.

5.1. Embodied Carbon

Renewables. The manufacturing (or embodied) carbon footprint for wind turbines ranges from 10-15 grams of CO_2 per kWh whereas the footprint for solar farms ranges from 40-70 grams of CO_2 per kWh [25]. These numbers are derived from a life cycle analysis and accounts for manufacturing costs and the expected amount of energy generated over the asset's lifetime. The average lifetime for solar panels is 25-30 years and that for wind turbines is 20 years.

Batteries. The manufacturing footprint of lithium-ion batteries ranges from 74 to 134 kilograms of CO_2 per KWh of battery capacity [15, 56]. The footprint includes material production, cell production and assembly, as well as end-of-life processing for the batteries, which is a necessary and challenging task [74]. The lifetime of the battery is calculated in terms of the number of discharge cycles. Utility-scale batteries, such as Tesla's Powerpack, last 3000-4000 cycles [59]. In this study, we assume one discharge cycle is used per day and hence battery lifetime is approximately ten years.

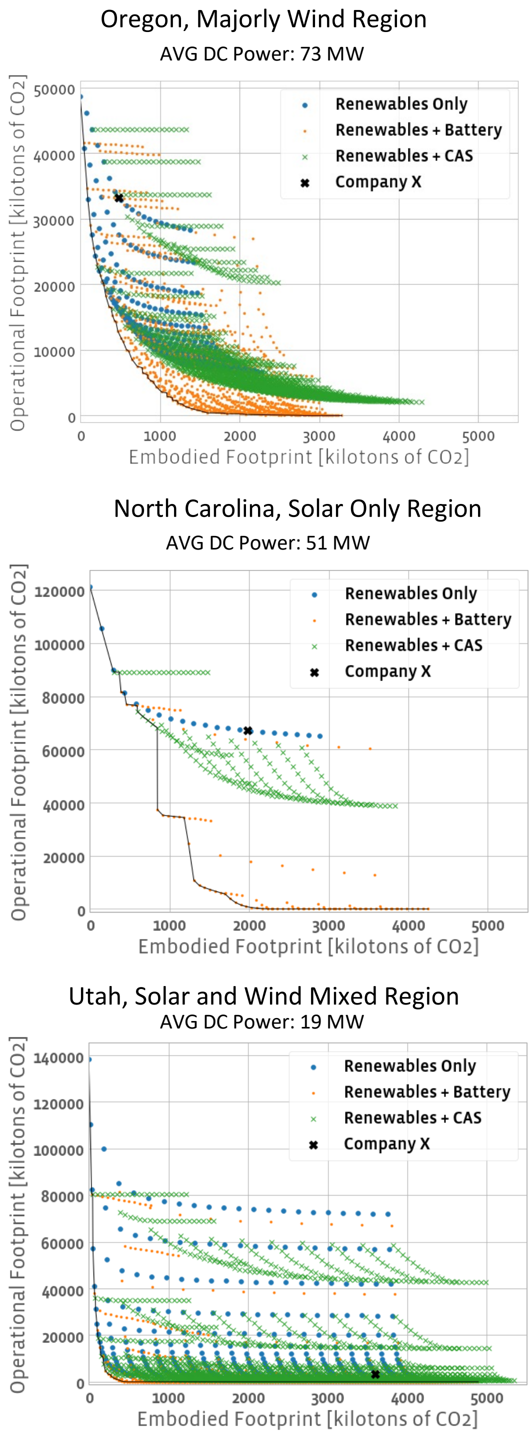

Figure 16: Operational and embodied footprint of the three solutions. Pareto frontier shows how the long tail to reach 100% renewable coverage can be shortened with complementary solutions.

Servers. The manufacturing footprint of servers is estimated to be 744.5 kg eq CO_2 [26], using an HPE ProLiant DL360 Gen10 server as a proxy. This server includes a single-socket CPU with 48 GB DRAM and has Thermal Design Power (TDP) of 85 Watts. Carbon measurements include its mainboard, SSD, daughterboard, enclosure, fans, transport and assembly. We estimate server lifetime of five years.

5.2. A Holistic Analysis

Figure 16 illustrates unique trade-offs between decreases in operational carbon (y-axis) and increases in embodied carbon (x-axis). In this evaluation, we assume 40% of datacenter workloads are delay-tolerant, a realistic flexible workload ratio [64], and can be deferred for carbon-aware scheduling. We examine three strategies: renewable energy generation alone, renewables with batteries, and renewables with carbon-aware scheduling (Section 4).

The space includes solutions that can significantly reduce a datacenter's overall carbon footprint. Reductions in operational carbon are an order of magnitude greater than increases in embodied carbon — $O(10^4)$ reductions versus $O(10^3)$ increases in kilotons of CO_2 . Yet datacenter operators must be careful in its pursuit of 24/7 coverage because some solutions incur much higher embodied carbon costs than others. Renewable generation alone is insufficient and solutions that combine renewable energy generation with batteries and scheduling are essential.

Batteries are essential and particularly cost effective. The Pareto frontier indicates that any solution for 24/7 carbon-free operations (i.e., zero operational carbon) must include renewable energy and batteries. Moreover, as 24/7 coverage increases, solutions that include batteries will incur smaller embodied carbon costs than solutions that rely solely on renewable energy and/or deploying additional servers to support carbon-aware scheduling.

Unfortunately, the Pareto frontier exhibits a long tail, which indicates increasingly expensive solutions required to reach full 24/7 coverage. For example, in Oregon, investments in renewable energy and batteries can quickly reduce operational CO_2 from 50M to 10M tons. But eliminating the last 10M tons of carbon will require significantly larger batteries.

Figure 17 details the most effective strategy for 24/7 carbon-free datacenter operation by geographic location and availability of renewable energy. We show the total carbon footprint, breaking down operational (solid) and embodied (cross-pattern) components. The carbon footprint is normalized relative to datacenter sizes, measured in MW of power capacity. We annotate each bar, identifying solutions that achieve full 100% 24/7 coverage (green stars) and those that make partial progress (red percentages).

Renewables Only. Relying solely on renewable energy generation incurs the highest embodied carbon costs in every geographic region. The 24/7 renewable coverage

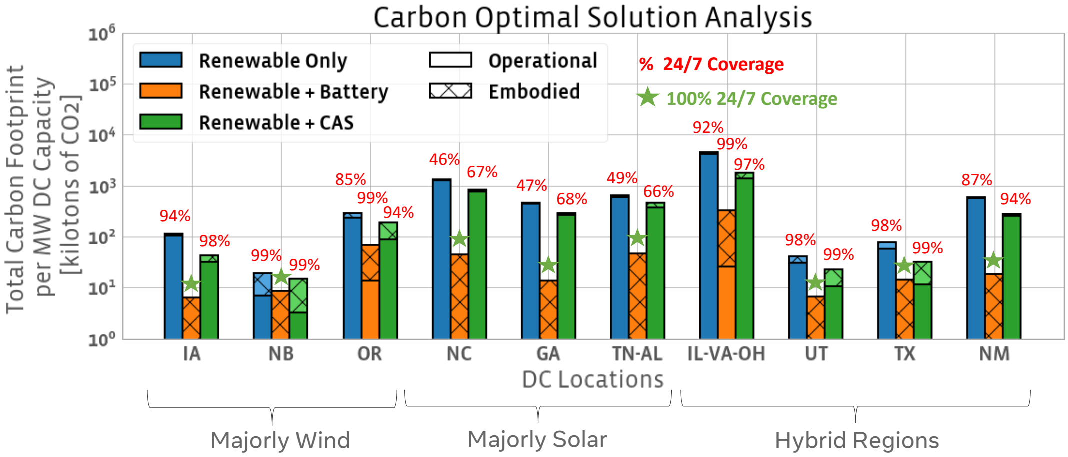

Figure 17: Plot shows the total footprint of the carbon-optimal setting of each solution per MW DC capacity broken down to operational and embodied components.

with renewables only ranges from 46% to 99% depending on geographical regions. Because renewable energy generation is intermittent, datacenter operators would need a large number of solar or wind farms to ensure sufficient supply during supply valleys. Even with significant investment, however, renewable energy supply fundamentally depends on weather and time of day, which leads to incomplete 24/7 coverage and higher operational carbon footprints that dominate the datacenter's total carbon footprint.

Figure 17 indicates the most effective renewable solutions include wind farms. A combination of wind and solar farms provides complementary generating assets and mitigates supply variance. Hybrid geographic regions, which use both wind and solar, achieve higher 24/7 coverage that ranges from 87% to 98%.

Regions that rely primarily on wind can also achieve high 24/7 coverage with careful datacenter site selection. Out of the thirteen locations, datacenters in Nebraska (a majorly-wind region), Utah and Texas (wind and solar hybrid regions) stand out as the best locations to minimize total carbon footprint and achieve highest coverage. Valleys in energy supply are shallower in these windy regions compared to those in others. In contrast, regions that rely primarily on solar (i.e., NC, GA, TN, AL) struggle to achieve full 24/7 coverage and incur the highest carbon footprints since solar energy is only available during parts of the day.

Renewables + Battery. The addition of batteries reduces the total carbon footprint by an order of magnitude in all geographic regions. The reduction is most pronounced in regions that rely only on solar energy. For nine of the thirteen datacenter regions, the most carbon-efficient solution deploys enough battery capacity to achieve 100% 24/7 coverage and completely eliminate the datacenter's operational carbon footprint. For the other four datacenter regions (OR, IL, VA, OH) 99% renewable energy coverage is the carbon-optimal solution. The battery capacity amount to achieve the optimal footprint ranges from 10MWh - 2000MWh for different datacenters and it would represent a utility scale battery capacity. Given hyperscale datacenters cost billions of dollars [51], this battery investment represents a small fraction of a data center's overall cost at current battery prices of $350/kWh [11].

Renewables + CAS. Carbon-aware scheduling provides an alternative to batteries, increasing 24/7 coverage by 1% to 21% across geographic regions. Carbon Explorer finds that deploying 6% to 76% additional server capacity allows the scheduler to move computation from periods when renewable energy is scarce to periods when it is abundant, thereby reducing the datacenter's overall carbon footprint. However, carbon-aware scheduling is constrained by the degree of workload flexibility and the number of provisioned servers available to process deferred jobs. Due to these constraints, scheduling alone is insufficient for full 24/7 coverage in regions characterized by many days with near zero renewable energy (e.g., wind in Oregon) or regions that rely exclusively on solar energy.

In summary, 24/7 coverage depends on renewable energy generation characteristics of the datacenter region. In several cases, achieving complete 24/7 coverage is neither feasible nor the most carbon-efficient solution. Solutions that complement investments in renewable energy are necessary. The addition of batteries or carbon-aware scheduling can reduce a datacenter's total carbon footprint by an order of magnitude. A combination of battery deployments and carbon-aware scheduling may offer additional improvement. How to make these decisions optimally is an open research question left for future work.

6. Discussion and Related Work

Renewable Energy. Prior academic research emphasizes renewable energy on-site at the datacenter [38]. Computation uses local, solar energy and minimizes energy consumed from the grid [18, 19]. The datacenter's power infrastructure is enhanced to switch between multiple types of local generators and microgrids [38, 39, 45]. And strategies are developed to deploy scale out servers and renewable generators in a modular fashion [35, 37]. These strategies seem sensible for edge and fog servers [36]. However, the hyperscale datacenters we study avoid many of these challenges. They do not need to manage local power generation because they have invested in renewable generation on the grid at scales that are unlikely on-site. Yet they improve sustainability through power purchase agreements. We study renewable energy across geographic locations and coordinate their installation with battery and server provisioning at scale.

Energy Storage. Batteries ensure datacenter availability but can also modulate the datacenter's demands for grid power [22, 23]. For datacenters that use renewable energy, batteries can mitigate intermittent supplies of solar and wind [41, 44]. Performance and efficiency vary with battery technology, motivating heterogeneous solutions [43, 46]. Battery aging can be mitigated by managing charge-discharge cycles and demand for stored energy [42]. We quantify energy storage required for 24/7 carbon-free computing and, without loss of generality, consider lithium-ion batteries for their attractive downward cost trajectory and acceptable ten-plus year lifetimes under simulated usage.

Battery technologies will impact data center design and management. The price of lithium-ion batteries is falling significantly, declining by 80% from 2015 to 2020 [52, 53]. These batteries have been deployed at scale and, for example, can supply 28 MW for four hours. Such operational parameters align with hyperscale datacenters, which are provisioned for 20 to 40 MW. Four hours of battery operation could significantly reduce demand response requirements from job scheduling.

Although our paper makes the case for energy storage, it does not explicitly prescribe an implementation strategy. Datacenter operators could collaborate with utility providers to invest in batteries on the grid just as they do for wind and solar farms. Alternatively, they could deploy batteries on-site at the datacenter. Datacenter may wish to implement custom battery charge-discharge policies, which have previously been explored at much smaller scales for uninterruptible power supplies [22, 23]. Whether these policies can be implemented in the form of contracts with grid operators is to be determined.

Finally, there are potentially environmental and health risks associated with the disposal of batteries. Spent LIBs contain toxic materials including heavy metals and flammable electrolytes, and therefore they need to be properly recycled and disposed in order not to cause contamination of the soil, water and air [55, 75]. This is another aspect that needs consideration when making large-scale battery deployments.

Carbon-Aware Scheduling. Time-series analysis accurately forecasts renewable supplies and datacenter demands for energy. Forecasts permit optimizing schedules of flexible jobs in response to energy supply [76]. Optimization objectives have accounted for electricity prices [47], carbon prices in cap-and-trade markets [34], the carbon-intensity of grid energy [54], and service quality [30]. Timely energy data is necessary for intelligent scheduling [4, 70]. We perform offline analyses to defer flexible computation and explore the design space for 24/7 carbon-free computing. A future implementation would benefit from prior schedulers.

Power Transmission. In addition to energy storage technologies, transmitting electricity generated from renewables is another potentially viable option. The recent technology breakthroughs in renewable energy generation, energy storage, and electricity transmission are energizing novel infrastructure development and deployment [72], such as, transmitting solar power from Australia underneath an ocean to Singapore [57]. High-voltage direct current (HVDC) transmission technology is becoming an attractive option to transmit electricity between renewable generation sites and power grids and for trans-national grids [60, 69]. While the HVDC technology has seen significant efficiency and energy capacity transmission improvement, it is currently less efficient as compared to data/computation scheduling across data centers in different geographic locations (by roughly two orders of magnitude). Thus, *locality* — from the perspectives of where energy is generated and where data resides — is important to keep in mind for the near future. As power transmission technologies becomes increasingly cost-effective, renewable energy availability will become more geographically inclusive. The impact of the power transmission dimension can be taken into the Carbon Explorer design space to understand the dynamics.

7. Conclusion

This paper presents *Carbon Explorer* — a design space exploration tool to enable carbon-optimal investment strategies. Carbon Explorer determines carbon-optimal settings across the dimensions of *investments on various renewable energy types*, *the amount of energy storage*, and *carbon-aware computation shifting* by considering geographically-dependent renewable energy availability characteristics and computation demand patterns at the data center scale. Carbon Explorer demonstrates that, depending on graphical locations, carbon-optimal strategies vary and that when embodied carbon footprint is considered, 100% 24/7 operational carbon-free computing may not always be carbon-optimal. We hope Carbon Explorer can guide future sustainability investments to achieve operational and embodied carbon footprint optimality.

Acknowledgements

We would like to thank Hsien-Hsin Sean Lee, Dimitrios Skarlatos, Max Balandat, Newsha Ardalani, Hugh Leather, Sal Candido for discussions and feedback on this work. We also thank Kim Hazelwood for supporting this work.

References

[1] *High Power Lithium Ion APR18650M1A*, A123 Systems, 2009, liFePO4 cell specifications.

[2] A. S. Andrae and T. Edler, "On global electricity usage of communication technology: trends to 2030," *Challenges*, vol. 6, no. 1, pp. 117–157, 2015.

[3] L. A. Barroso and U. Hölzle, "The case for energy-proportional computing," *Computer*, vol. 40, no. 12, pp. 33–37, 2007.

[4] N. Bashir, T. Guo, M. Hajiesmaili, D. Irwin, P. Shenoy, R. Sitaraman, A. Souza, and A. Wierman, "Enabling sustainable clouds: The case for virtualizing the energy system," *arXiv preprint arXiv:2106.08872*, 2021.

[5] California ISO, "Managing oversupply," http://www.caiso.com/informed/Pages/ManagingOversupply.aspx, 2021.

[6] California State Commission, "2020 total system electric generation," https://www.energy.ca.gov/data-reports/energy-almanac/california-electricity-data/2020-total-system-electric-generation, 2020.

[7] J. Chase, "Demand response for computing centers," 2010.

[8] A. Chien, "Characterizing opportunity power in the california independent system operator (caiso) in years 2015-2017," *Energy and Earth Science*, 2020.

[9] A. A. Chien, F. Yang, and C. Zhang, "Characterizing curtailed and uneconomic renewable power in the mid-continent independent system operator," *AIMS Energy*, vol. 6, no. 2, pp. 376–401, 2018.

[10] Citation anonymised for blind review, "xxx," xxx, 2021.

[11] W. Cole, A. W. Frazier, and C. Augustine, "Cost projections for utility-scale battery storage: 2021 update," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2021.

[12] B. Diouf and R. Pode, "Potential of lithium-ion batteries in renewable energy," *Renewable Energy*, vol. 76, pp. 375–380, 2015.

[13] DOE Global Energy Storage Database, "Vistra moss landing energy storage," https://sandia.gov/ess-ssl/gesdb/public/projects.html#1331, 2022.

[14] E. Ela, "Using economics to determine the efficient curtailment of wind energy," National Renewable Energy Lab.(NREL), Golden, CO (United States), Tech. Rep., 2009.

[15] E. Emilsson and L. Dahllöf, "Lithium-ion vehicle battery production-status 2019 on energy use, co2 emissions, use of metals, products environmental footprint, and recycling," 2019.

[16] Facebook, "Advancing renewable energy through green tariffs," https://sustainability.fb.com/wp-content/uploads/2020/12/FB_Green-Tariffs.pdf, 2021.

[17] Facebook, "How we're helping fight climate change," https://about.fb.com/news/2021/06/2020-sustainability-report-how-were-helping-fight-climate-change/, 2021.

[18] Í. Goiri, W. Katsak, K. Le, T. D. Nguyen, and R. Bianchini, "Parasol and greenswitch: Managing datacenters powered by renewable energy," *ACM SIGPLAN Notices*, vol. 48, no. 4, pp. 51–64, 2013.

[19] Í. Goiri, W. Katsak, K. Le, T. D. Nguyen, and R. Bianchini, "Designing and managing data centers powered by renewable energy," *IEEE Micro*, vol. 34, no. 3, pp. 8–16, 2014.

[20] Google, "Google's green ppas: What, how and why," https://static.googleusercontent.com/media/www.google.com/en//green/pdfs/renewable-energy.pdf, 2013.

[21] Google, "24/7 by 2030: Realizing a carbon-free future," 2020. [Online]. Available: https://www.gstatic.com/gumdrop/sustainability/247-carbon-free-energy.pdf

[22] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, "Benefits and limitations of tapping into stored energy for datacenters," in *2011 38th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2011, pp. 341–351.

[23] S. Govindan, D. Wang, A. Sivasubramaniam, and B. Urgaonkar, "Leveraging stored energy for handling power emergencies in aggressively provisioned datacenters," in *Proceedings of the seventeenth international conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, 2012, pp. 75–86.

[24] U. Gupta, Y. Kim, S. Lee, J. Tse, H. S. Lee, G. Wei, D. Brooks, and C. Wu, "Chasing carbon: The elusive environmental footprint of computing," in *Proceedings of the International Symposium on High-Performance Computer Architecture (HPCA)*, 2021.

[25] G. Heath, "Life cycle assessment harmonization," https://www.nrel.gov/analysis/life-cycle-assessment.html, 2021.

[26] Hewlett Packard Enterprise, "Hpe product carbon footprint, hpe proliant dl360 gen10 server," https://www.hpe.com/psnow/doc/a50002430enw?jumpid=in_lit-psnow-red, 2021.

[27] N. Jones, "How to stop data centres from gobbling up the world's electricity," *Nature*, vol. 561, no. 7722, pp. 163–167, 2018.

[28] L. Joppa, "Made to measure: Sustainability commitment progress and updates," https://blogs.microsoft.com/blog/2021/07/14/made-to-measure-sustainability-commitment-progress-and-updates/, 2021.

[29] Justine Calma, "Microsoft is changing the way it buys renewable energy," https://www.theverge.com/2021/7/14/22574431/microsoft-renewable-energy-purchases, 2021.

[30] W. Katsak, Í. Goiri, R. Bianchini, and T. D. Nguyen, "Greencassandra: Using renewable energy in distributed structured storage systems," in *2015 Sixth International Green and Sustainable Computing Conference (IGSC)*. IEEE, 2015, pp. 1–8.

[31] F. Kazhamiaka, C. Rosenberg, and S. Keshav, "Tractable lithium-ion storage models for optimizing energy systems," *Energy Informatics*, vol. 2, no. 1, pp. 1–22, 2019.

[32] R. Koningstein, "We now do more computing where there's cleaner energy," https://blog.google/outreach-initiatives/sustainability/carbon-aware-computing-location/, 2021.

[33] V. Kontorinis, L. E. Zhang, B. Aksanli, J. Sampson, H. Homayoun, E. Pettis, D. M. Tullsen, and T. Simunic Rosing, "Managing distributed ups energy for effective power capping in data centers," in *Proceedings of the International Symposium on Computer Architecture (ISCA)*, 2012.

[34] K. Le, O. Bilgir, R. Bianchini, M. Martonosi, and T. D. Nguyen, "Managing the cost, energy consumption, and carbon footprint of internet services," *ACM SIGMETRICS Performance Evaluation Review*, vol. 38, no. 1, pp. 357–358, 2010.

[35] C. Li, Y. Hu, J. Gu, J. Yuan, and T. Li, "Oasis: scaling out datacenter sustainably and economically," *IEEE Transactions on Parallel and Distributed Systems*, 2016.

[36] C. Li, Y. Hu, L. Liu, J. Gu, M. Song, X. Liang, J. Yuan, and T. Li, "Towards sustainable in-situ server systems in the big data era," in *2015 ACM/IEEE 42nd Annual International Symposium on Computer Architecture (ISCA)*, 2015, pp. 14–26.

[37] C. Li, Y. Hu, R. Zhou, M. Liu, L. Liu, J. Yuan, and T. Li, "Enabling datacenter servers to scale out economically and sustainably," in *Proceedings of the 46th annual IEEE/ACM international symposium on microarchitecture*, 2013, pp. 322–333.

[38] C. Li, A. Qouneh, and T. Li, "iSwitch: Coordinating and optimizing renewable energy powered server clusters," *ACM SIGARCH Computer Architecture News*, vol. 40, no. 3, pp. 512–523, 2012.

[39] C. Li, R. Zhou, and T. Li, "Enabling distributed generation powered sustainable high-performance data center," in *2013 IEEE 19th International Symposium on High Performance Computer Architecture (HPCA)*. IEEE, 2013, pp. 35–46.

[40] L. Lin and A. Chien, "Characterizing stranded power in the ERCOT in years 2012-2019: A preliminary report," *University of Chicago CS Tech Report*, 2020.

[41] L. Liu, H. Sun, C. Li, Y. Hu, J. Xin, N. Zheng, and T. Li, "RE-UPS: An adaptive distributed energy storage system for dynamically managing solar energy in green datacenters," *the Journal of Supercomputing*, vol. 72, no. 1, pp. 295–316, 2016.

[42] L. Liu, C. Li, H. Sun, Y. Hu, J. Gu, and T. Li, "Baat: Towards dynamically managing battery aging in green datacenters," in *2015 45th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*. IEEE, 2015, pp. 307–318.

[43] L. Liu, C. Li, H. Sun, Y. Hu, J. Gu, T. Li, J. Xin, and N. Zheng, "Heb: Deploying and managing hybrid energy buffers for improving datacenter efficiency and economy," in *2015 ACM/IEEE 42nd Annual International Symposium on Computer Architecture (ISCA)*, 2015, pp. 463–475.

[44] L. Liu, H. Sun, Y. Hu, J. Xin, N. Zheng, and T. Li, "Leveraging distributed ups energy for managing solar energy powered data centers," in *International Green Computing Conference*. IEEE, 2014, pp. 1–8.

[45] L. Liu, H. Sun, C. Li, Y. Hu, N. Zheng, and T. Li, "Towards an adaptive multi-power-source datacenter," in *Proceedings of the 2016 International Conference on Supercomputing*, 2016, pp. 1–11.

[46] L. Liu, H. Sun, C. Li, T. Li, J. Xin, and N. Zheng, "Exploring highly dependable and efficient datacenter power system using hybrid and hierarchical energy buffers," *IEEE Transactions on Sustainable Computing*, 2018.

[47] Z. Liu, I. Liu, S. Low, and A. Wierman, "Pricing data center demand response," *ACM SIGMETRICS Performance Evaluation Review*, vol. 42, no. 1, pp. 111–123, 2014.

[48] S. Malla, Q. Deng, Z. Ebrahimzadeh, J. Gasperetti, S. Jain, P. Kondety, T. Ortiz, and D. Vieira, "Coordinated priority-aware charging of distributed batteries in oversubscribed data centers," in *Proceedings of the International Symposium on Microarchitecture (ISCA)*, 2020.

[49] S. Malla, Q. Deng, Z. Ebrahimzadeh, J. Gasperetti, S. Jain, P. Kondety, T. Ortiz, and D. Vieira, "Coordinated priority-aware charging of distributed batteries in oversubscribed data centers," in *2020 53rd Annual IEEE/ACM International Symposium on Microarchitecture (MICRO)*. IEEE, 2020, pp. 839–851.

[50] E. Masanet, A. Shehabi, N. Lei, S. Smith, and J. Koomey, "Recalibrating global data center energy-use estimates," *Science*, vol. 367, no. 6481, pp. 984–986, 2020.

[51] R. Miller, "Google has invested $5 billion in its iowa data centers," https://datacenterfrontier.com/google-has-invested-5-billion-in-its-iowa-data-centers/, 2022.

[52] K. Mongird, V. V. Viswanathan, P. J. Balducci, M. J. E. Alam, V. Fotedar, V. S. Koritarov, and B. Hadjerioua, "Energy storage technology and cost characterization report," Pacific Northwest National Lab.(PNNL), Richland, WA (United States), Tech. Rep., 2019.

[53] National Renewable Energy Laboratory, "Declining renewable costs drive focus on energy storage," 2020. [Online]. Available: https://www.nrel.gov/news/features/2020/declining-renewable-costs-drive-focus-on-energy-storage.html

[54] A. Radovanovic, R. Koningstein, I. Schneider, B. Chen, A. Duarte, B. Roy, D. Xiao, M. Haridasan, P. Hung, N. Care *et al.*, "Carbon-aware computing for datacenters," *arXiv preprint arXiv:2106.11750*, 2021.

[55] M. A. Rajaeifar, P. Ghadimi, M. Raugei, Y. Wu, and O. Heidrich, "Challenges and recent developments in supply and value chains of electric vehicle batteries: A sustainability perspective," *Resources, Conservation and Recycling*, vol. 180, p. 106144, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0921344921007527

[56] R. Rapier, "Estimating the carbon footprint of utility-scale battery storage," https://www.forbes.com/sites/rrapier/2020/02/16/estimating-the-carbon-footprint-of-utility-scale-battery-storage/?sh=167c8bd97adb, 2020.

[57] R. Rapier, "Transmitting solar power underneath an ocean," 2020. [Online]. Available: https://4thgeneration.energy/transmitting-solar-power-underneath-an-ocean/

[58] H. Ritchie and M. Roser, "Energy," *Our World in Data*, 2020, https://ourworldindata.org/energy.

[59] Z. Shahan, "Tesla Megapack, Powerpack, & Powerwall Battery Storage Prices Per KWh," https://cleantechnica.com/2020/10/05/tesla-megapack-powerpack-powerwall-battery-storage-prices/, 2020.

[60] Siemens, "Fact sheet: High-voltage direct current transmission (hvdc)," 2014. [Online]. Available: https://assets.new.siemens.com/siemens/assets/api/uuid:d5c5f4ae-d9f6-49e9-b68b-85bb7ceb4f41/factsheet-hvdc-e.pdf

[61] C. Tang, K. Yu, K. Veeraraghavan, J. Kaldor, S. Michelson, T. Kooburat, A. Anbudurai, M. Clark, K. Gogia, L. Cheng, B. Christensen, A. Gartrell, M. Khutornenko, S. Kulkarni, M. Pawlowski, T. Pelkonen, A. Rodrigues, R. Tibrewal, V. Venkatesan, and P. Zhang, "Twine: A unified cluster management system for shared infrastructure," in *Proceedings of the USENIX Symposium on Operating Systems Design and Implementation*, 2020.

[62] The U.S. Energy Information Administration, "Renewable energy explained," https://www.eia.gov/energyexplained/renewable-sources/, 2021.

[63] The White House, "FACT SHEET: President Biden Renews U.S. Leadership on World Stage at U.N. Climate Conference (COP26)," https://www.whitehouse.gov/briefing-room/statements-releases/2021/11/01/fact-sheet-president-biden-renews-u-s-leadership-on-world-stage-at-u-n-climate-confe 2021.

[64] M. Tirmazi, A. Barker, N. Deng, M. E. Haque, Z. G. Qin, S. Hand, M. Harchol-Balter, and J. Wilkes, "Borg: the next generation," in *Proceedings of the Fifteenth European Conference on Computer Systems*, 2020, pp. 1–14.

[65] United Nations, "24/7 carbon-free energy compact," https://www.un.org/en/energy-compacts/page/compact-247-carbon-free-energy.

[66] U.S. Department of Defense (DoD) and the U.S. General Services Administration (GSA), "Dod, gsa announce rfi to gather information for supplying 24/7 carbon pollution-free electricity for federal government," https://www.defense.gov/News/Releases/Release/Article/2921646/dod-gsa-announce-rfi-to-gather-information-for-supplying-247-carbon-pollution-f/.

[67] U.S. Energy Information Administration, "Annual energy outlook," https://www.eia.gov/outlooks/archive/aeo21/, 2021.

[68] U.S. Energy Information Administration, "Hourly electric grid monitor," https://www.eia.gov/electricity/gridmonitor/dashboard/electric_overview/US48/US48, 2021.

[69] K. Vaillancourt, "Electricity transmission and distribution," 2014. [Online]. Available: https://iea-etsap.org/E-TechDS/PDF/E12_el-t&d_KV_Apr2014_GSOK.pdf

[70] A. Wierman, Z. Liu, I. Liu, and H. Mohsenian-Rad, "Opportunities and challenges for data center demand response," in *International Green Computing Conference*. IEEE, 2014, pp. 1–10.

[71] P. Wiesner, I. Behnke, D. Scheinert, K. Gontarska, and L. Thamsen, "Let's wait awhile: How temporal workload shifting can reduce carbon emissions in the cloud," *arXiv preprint arXiv:2110.13234*, 2021.

[72] C.-J. Wu, S. Manne, P. Ranganathan, S. Bird, and S. Greenstein, "Socio-technological challenges and opportunities: Paths forward," 2021.

[73] C.-J. Wu, R. Raghavendra, U. Gupta, B. Acun, N. Ardalani, K. Maeng, G. Chang, F. A. Behram, J. Huang, C. Bai, M. Gschwind, A. Gupta, M. Ott, A. Melnikov, S. Candido, D. Brooks, G. Chauhan, B. Lee, H.-H. S. Lee, B. Akyildiz, M. Balandat, J. Spisak, R. Jain, M. Rabbat, and K. Hazelwood, "Sustainable ai: Environmental implications, challenges and opportunities," 2021.

[74] Y. Yang, E. G. Okonkwo, G. Huang, S. Xu, W. Sun, and Y. He, "On the sustainability of lithium ion battery industry–a review and perspective," *Energy Storage Materials*, 2020.

[75] X. Zeng, J. Li, and L. Liu, "Solving spent lithium-ion battery problems in china: Opportunities and challenges," *Renewable and Sustainable Energy Reviews*, vol. 52, pp. 1759–1767, 2015.

[76] C. Zhang and A. A. Chien, "Scheduling challenges for variable capacity resources," in *Workshop on Job Scheduling Strategies for Parallel Processing (JSSPP)*. Springer, 2021, pp. 190–209.

[77] C. Zhang, A. G. Kumbhare, I. Manousakis, D. Zhang, P. A. Misra, R. Assis, K. Woolcock, N. Mahalingam, B. Warrier, D. Gauthier *et al.*, "Flex: High-availability datacenters with zero reserved power," in *2021 ACM/IEEE 48th Annual International Symposium on Computer Architecture (ISCA)*. IEEE, 2021, pp. 319–332.

[78] J. Zheng, A. A. Chien, and S. Suh, "Mitigating curtailment and carbon emissions through load migration between data centers," *Joule*, vol. 4, no. 10, pp. 2208–2222, 2020.