Adapting Datacenter Capacity for Greener Datacenters and Grid

Liuzixuan Lin
University of Chicago
Chicago, IL, USA
lzixuan@uchicago.edu

Andrew A. Chien
University of Chicago & Argonne National Lab
Chicago, IL, USA
aachien@uchicago.edu

ABSTRACT

Cloud providers are adapting datacenter (DC) capacity to reduce carbon emissions. With hyperscale datacenters exceeding 100 MW individually, and in some grids exceeding 15% of power load, DC adaptation is large enough to harm power grid dynamics, increasing carbon emissions, power prices, or reduce grid reliability.

To avoid harm, we explore coordination of DC capacity change varying scope in space and time. In space, coordination scope spans a single datacenter, a group of datacenters, and datacenters with the grid. In time, scope ranges from online to day-ahead. We also consider what DC and grid information is used (e.g. real-time and day-ahead average carbon, power price, and compute backlog). For example, in our proposed PlanShare scheme, each datacenter uses day-ahead information to create a capacity plan and shares it, allowing global grid optimization (over all loads, over entire day).

We evaluate DC carbon emissions reduction. Results show that local coordination scope fails to reduce carbon emissions significantly (3.2%–5.4% reduction). Expanding coordination scope to a set of datacenters improves slightly (4.9%–7.3%). PlanShare, with grid-wide coordination and full-day capacity planning, performs the best. PlanShare reduces DC emissions by 11.6%–12.6%, 1.56x–1.26x better than the best local, online approach's results. PlanShare also achieves lower cost. We expect these advantages to increase as renewable generation in power grids increases. Further, a known full-day DC capacity plan provides a stable target for DC resource management.

CCS CONCEPTS

• **Applied computing → Data centers**; • **Hardware → Power and energy**; **Impact on the environment**.

KEYWORDS

Data centers, Carbon emissions, Capacity adaptation, Power management, Adaptive loads

ACM Reference Format:
Liuzixuan Lin and Andrew A. Chien. 2023. Adapting Datacenter Capacity for Greener Datacenters and Grid. In *The 14th ACM International Conference on Future Energy Systems (e-Energy '23), June 20–23, 2023, Orlando, FL, USA.* ACM, New York, NY, USA, 14 pages. https://doi.org/10.1145/3575813.3595197

e-Energy '23, June 20–23, 2023, Orlando, FL, USA

ACM ISBN 979-8-4007-0032-3/23/06.
https://doi.org/10.1145/3575813.3595197

1 INTRODUCTION

With the commercial success of internet-scale applications and cloud computing, cloud infrastructure has grown rapidly. A recent article documented the addition of over 50 datacenters a year by a single cloud provider [79]. By revenue, Amazon, Microsoft, and Google's cloud growth rates have exceeded 30% annually for the past 5 years [30, 46, 74]. The accelerated digitalization since COVID-19 [73], and an accelerating adoption of machine learning (aka artificial intelligence) are both driving an acceleration of datacenter growth [80, 88]. In 2021, the power consumption of these three cloud providers exceeded 62 TWh [8, 36, 46], equivalent to the power consumed by 6.2 million American homes. In 2022 they purchased 14 GW of renewable generation capacity, but not enough to offset their power use [54]. Some estimates project that datacenter power consumption will grow to 10% of global electricity use by 2030 [45, 56, 65]. Today's largest hyperscaler sites are multiple buildings with total power of 200 MW to 1 GW [3, 9, 34, 69, 79].

In many power grids, datacenters are already major load contributors. In Virginia, datacenters account for 12% of power consumption (2022), and will reach 18% in 2027 and 22% in 2032 [28, 29]. In Ireland, datacenters account for 14% of national electricity use (2022) [13] and may be 30% by 2029 [31]. With continued cloud and artificial intelligence growth, datacenters are expected to exceed 10% or even 20% of load in many power grids [41, 45, 65].

Rapid computing growth raises concerns about carbon emissions [3]. Cloud providers often purchase renewable power (long-term contracts) or renewable offsets (renewable energy credits—RECs) to "offset" their power use [19, 74]. However, these contracts are accounting arrangements, not actual power transfers. Despite such arrangements and even full offsetting, cloud datacenters consume large quantities of fossil-fuel generated power [3, 10]. Worse, the growth of datacenters can threaten grid stability, blocking DC projects in Ireland and Northern Virginia's grids. [47, 48].

Adapting capacity (temporal shifting) to reduce DC carbon emissions has been explored for over a decade [23, 24, 32, 60], and to reduce power cost [7, 57, 62, 63, 81, 87, 93]. For example, cloud providers' "24×7" goals to match hourly power use and renewable generation typically involve temporal shifting [4, 19, 33, 35, 78]. These approaches all seek to align compute load with low-carbon or low-price power, using online control techniques (see Figure 1). Such efforts along with those we propose in this paper are increasingly crucial as power grids around the world driven by aggressive public policy are rapidly decarbonizing by adding new renewable generation [12, 25, 75, 86].

DC shifting efforts typically focus on internal challenges of DC resource management. A recent system forces shifting by creating a daily compute capacity plan to enable compute resource management [78]. However, most shifting efforts ignore the negative

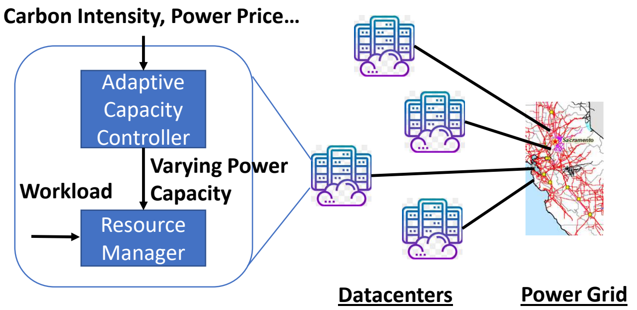


Figure 1: Datacenters can adapt power capacity to align compute load with periods of lower carbon emissions.

impacts of varying datacenter capacity on grid dynamics (e.g. triggering unnecessary generator starts and load shedding). With large datacenters, such impacts are critical. Load shifting schemes that are productive with small datacenters can be ineffective and even damage grid performance if used with large datacenters. In Section 3 (Problem), we show an example where such approaches create an 8% increase in datacenter carbon emissions due to overshifting. One consequence of this new insight is that there is no known solution to coordinate large-scale datacenter capacity adaptation and power grids. To explore this new problem, researchers have begun to study capacity adaptation coupled with power grid models [56, 58, 68].

To avoid the harm and effectively reduce DC carbon emissions with capacity adaptation, we explore coordination of DC capacity adaptation that varies in coordination scope—space and time. In space, we consider spans of a single datacenter, a group of datacenters, and datacenters with the power grid. In time, we consider online and day-ahead. We also vary datacenter and grid information used (e.g. real-time and day-ahead metrics, compute backlog). For example, in PlanShare, each datacenter uses day-ahead grid information to create a full-day capacity plan, and then shares it with the grid, so the grid can optimize generation and transmission with a known, but varying datacenter capacity schedule. Our evaluation reports datacenter carbon emissions reduction relative to the fixed DC capacity scenario, as well as cost impacts on both datacenters and other customers.

Specific contributions of the paper include:

- With local coordination scope, we use three grid metrics (average carbon intensity (ACI), grid price (Price), and locational marginal price (LMPrice)) for capacity adaptation. LMPrice is most effective, reducing datacenter carbon emissions 1–5% vs. ACI and 0.7–1.5% vs. Price. Exploiting hourly future price information (+2.3%) and step size (+2.3%) further improves local adaptation based on LMPrice, achieving datacenter carbon reduction of 10%.
- Expanding the coordination scope to a group of datacenters with a coordinator that limits aggregate behaviors gives only a small improvement over local adaptation, reducing datacenter carbon emissions by 7.3%.
- PlanShare, datacenter-grid coordination and full-day capacity planning, achieves greatest benefits, decreasing datacenter carbon emissions by 11.6%–12.6%. This grid-wide coordination is 1.56x–1.26x better than the best local, online approach. The key costs (grid dispatch cost, customer power cost) are also reduced. Further, for datacenters, the 24-hour capacity plan provides a stable target enabling more efficient compute resource management. PlanShare satisfies datacenter and power grid objectives.

The remainder of the paper includes Background (Section 2) and in Section 3 we describe the challenge of adapting large-datacenter capacity. In Section 4 we discuss datacenter capacity adaptation approaches. Section 5 introduces the methodology of grid-coupled simulation. In Section 6 we evaluate carbon reduction of the different capacity adaptation approaches, considering impacts on datacenter and non-datacenter grid customers. Finally, in Section 7 and 8, we discuss related work, summarize results, and discuss future research directions.

2 BACKGROUND

2.1 Growth of Cloud Datacenters in Power Grids

Hyperscale cloud providers (e.g. Amazon, Microsoft, Google, Alibaba) are building larger and more datacenters to meet the needs of digitalization, which accelerated since the COVID-19 pandemic [29, 73, 82]. The rapid growth of hyperscale cloud industry is reflected in its power consumption: 1% of worldwide power consumption in 2018, projections suggest it could exceed 10% by 2030 [45, 56, 65]. Similar rapid growth rates are also supported by corporate renewable power purchase data [46, 74].

As a result, datacenters have become the key driver of new electric power demand, direct cause of new emissions, as well as construction of power plants, transmission, and energy storage infrastructure [28, 47, 70]. In regions where datacenters account for large fraction of power consumption (e.g. Virginia—12% in 2022 and 22% in 10 years, Ireland—14% in 2022 and 30% by 2029), the challenges in grid operation and decarbonization have become evident. For example, Ireland's state grid operator is canceling datacenter projects [48] unless the datacenters bring their own renewable generation or adapt their power demand. Similar story is happening in more and more parts of the world as the scale and reach of datacenters continue to grow [3, 41, 45, 47, 65].

From the cloud provider perspective, such exponential growth in datacenter power consumption first translates into growing power cost, as power price [5] don't decrease at the same speed. Also, it raises concern about carbon emissions growth. Obscured by carbon offset or renewable purchase, additional datacenters can still result in more fossil fuel consumption if the load is not aligned with renewable generation [10, 27, 31]. To improve sustainability, Google and Microsoft seek to hourly match their datacenter power capacity with renewable generation (24×7, 100/100/0 [35, 46]), with attempts such as capacity adaptation [78]. While the current scale of load change is only a few percent of capacity [78], it is expected to increase in future (discussed in Section 5.2). Furthermore, such adaptation behaviors can disturb the grid even at a small percentage of capacity. There are already reports of power variations from supercomputers (all on or all off) affecting grid stability [83]. For gigawatt cloud datacenters, a 10% load change produces a 100 MW swing, similar to the dynamic range we study; a 40% load change, which is not unusual for some diurnal peak to trough, would be 400 MW!

2.2 Dynamics of a Renewable-based Power Grid

Aligned with the carbon reduction goals (e.g. halved by 2030 and net-zero by 2050 [66]) that seek to limit global warming, recent years have seen the rise of renewable sources in energy mix of many power grids across the world, such as 34% in California (2021), 27% in China (2020), and 22% in Europe (2020). There are more ambitious goals for this decade. For example, California aims at 60% renewable fraction by 2030 [22], and Germany plans to phase out coal power plants (30% of electricity supply in 2021) by 2030 [89].

Integrating intermittent renewable sources (mainly wind and solar) is challenging for the power grid. For example, renewable generation can be wasted due to temporal mismatch with energy demand and transmission limits, producing both negative-priced power (aka stranded power) [16, 43, 55] and even"curtailment" [11, 16, 18, 39, 40, 55] where the renewable generation is wasted (not used by the power grid). A complementary problem is generation shortage, such as under extreme weather (heat, storms) where wind or solar generation can be dramatically lower. One potential solution to this challenge is to increase supply or demand flexibility, and corresponding methods include energy storage and adaptive loads [43, 71].

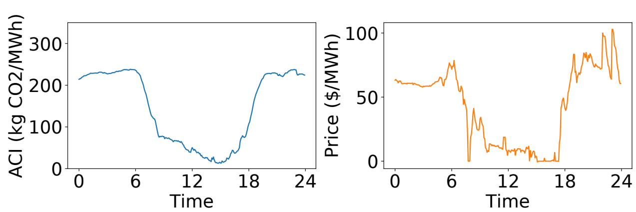


Figure 2: CAISO's Daily Average Carbon Intensity and Price Variation, 2022/05/02. Left: Average Carbon Intensity (kg CO2/MWh), Right: Grid Price ($/MWh). Source: CAISO.

The growth and intermittency of renewable generation produce time-varying grid metrics, such as carbon intensity (carbon emissions per MWh energy consumption) and power price. For example, in California's power grid, the average carbon intensity keeps low during the daytime when solar generation dominates, but it climbs up to about 200 kg CO_2/MWh as natural gas generators are up (Figure 2, left). In addition, as wind and solar generation is usually bid low due to zero fuel cost, the power price correlates with carbon intensity. The price can drop to near zero when there is excess renewable generation (Figure 2, right)! Adaptive loads such as datacenters, electric vehicles, and smart appliances may exploit such variation to reduce their carbon emissions or power cost.

3 PROBLEM

Cloud datacenters seek to reduce carbon emissions by aligning power use with plentiful renewable generation (low-carbon power). However, the carbon-intensity of grid power is determined by complex interaction of load, transmission, available generators and their ramping rate-limits. Grid dynamics can be opaque because power grids are a critical infrastructure and power markets have fierce economic competition. Together, these factors limit the real-time telemetry available to guide capacity adaptation. Therefore, it is challenging for large-scale cloud datacenters to dynamically choose capacities that reliably reduce carbon emissions. Often, load shifts achieve only a fraction of anticipated benefits, or worse, they can harm themselves or other customers by increasing prices or carbon emissions.

One example of this problem is "overshifting". Simulating a Spring day with our grid model (Figure 3), with intelligent local control based on grid *average carbon intensity (ACI)* [20, 64], all datacenters decrease capacity at hour 5 (4 am) and increase capacity at hour 15 (2 pm), causing the entire grid's ACI to increase by a total equivalent to 8% of datacenter daily carbon emissions (vs. fixed capacity)! Why is overshifting a problem? The local capacity control of the datacenters reacts to a shared metric in unison, increasing or decreasing capacity at the same time (2nd row). The large capacity changes, combined across datacenters, oversubscribe the opportunity. In order to maintain grid balance, the market responds by dispatching additional generation (fossil-fuel!), and thereby increasing carbon emissions.

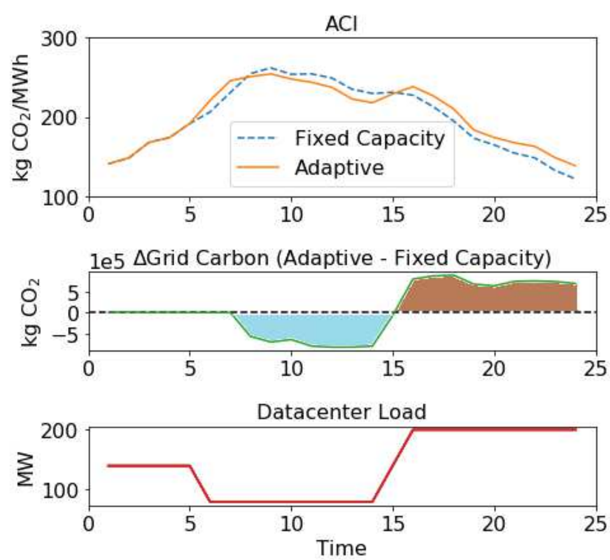


Figure 3: Overshifting: local, online adaptation using average carbon intensity (ACI) shown in the Top increases datacenter carbon emissions by 8% (Middle) because they all make the same capacity adaptation (Bottom).

Thus the challenge for datacenters is: **How to adapt capacity to reduce operational carbon emissions?** That is, how to align power use with low-carbon opportunity without disturbing the grid. To achieve this, capacity adaptation must:

(1) Identify opportunity: how to find the right times to increase and decrease capacity to reduce their carbon emissions?

(2) Avoid contention or overshifting: how to share the opportunity with other potential adaptive loads in the grid? (avoid oversubscribing the opportunity)
(3) Avoid harming others: how to ensure capacity adaptation does not harm others—load participants (consumers, other companies, other datacenters) by increasing prices or carbon emissions, or generators and grid resilience?

4 APPROACH

To address the challenges, we explore different datacenter capacity adaptation approaches and couple them to grid simulations to evaluate the **realized impacts**. The approaches vary in two dimensions of coordination scope: space and time. The space ranges from a single datacenter to a group and then to datacenters with the grid. The time includes online (real-time) to day-ahead (24 hours). In Figure 4, we illustrate three spatial scopes of coordination. We also vary the datacenter and grid information used for coordination.

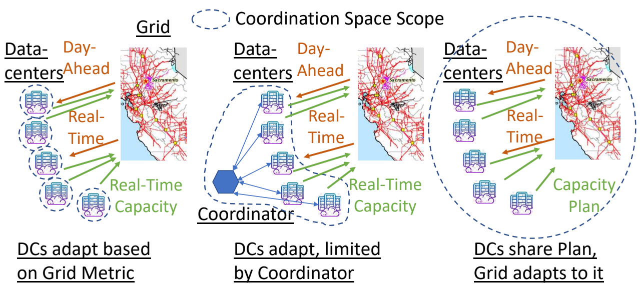


Figure 4: Adaptation Approaches vary in Coordination Scope.

Representing a classes of control techniques in previous work, **local** adaptation (Figure 4, left) makes hourly decisions using real-time and future metrics from the grid. Our studies of local adaptation show the limitations of these approaches. But first, we compare grid metrics (e.g. average carbon intensity, power price) that datacenters can use to drive local capacity adaptation.

At times, local adaptation can increase carbon emissions (e.g. overshifting in Figure 3). We consider use of an external **coordinator** that coordinates a group of datacenters to eliminate the harm (Figure 4, middle). In this scheme, each DC makes an hourly request to adapt capacity, and the external coordinator limits the group's total capacity change.

Finally, we consider a new approach that expands coordination scope to the entire grid. In this approach each DC **plans ahead**—making a binding capacity plan 24 hours in advance based on forecast grid information, and **shares the capacity plan with the grid**. The plan provides datacenter load certainty to the grid, enabling it to optimize generation and transmission scheduling under dynamic constraints that span single or many hours.

We introduce the algorithms for each approach, evaluating them in Section 6. We report DC carbon reduction with realistic levels of wind generation and datacenter load fractions (3.5%–14%), framing what could happen today or in the next 5–10 years. For clarity of exposition, we focus on a 30-DC scenario, representing about 10% of grid load that approximates the current levels in Northern Virginia (12%) or Ireland (14%), and vary the wind penetration (ratio of average wind generation to grid energy demand) from 15% to 60% (2015 level–2050 target). We also report how DC capacity adaptation creates grid impacts on power prices and how the impacts correlate with datacenter capacity change.

5 DATACENTER AND GRID-COUPLED SIMULATION METHODOLOGY

This section describes the framework for evaluating the datacenter capacity adaptation approaches. Datacenters adapt capacity to grid metrics (Section 5.1), respecting the capacity flexibility constraints (Section 5.2). Resulting time-varying loads affect the dynamics (e.g. pricing, generation) in the power grid (Section 5.3).

5.1 Grid Metrics for DC Capacity Adaptation

Power grids have complex dynamics, so the "best" grid metric[1] for reducing carbon emissions is an open research question [20, 58]. A good metric should enable carbon reduction and be available in most or all power grids. We consider several candidates:

- **Average carbon intensity or ACI (kg CO_2/MWh)** is the carbon emissions per MWh energy consumption in the grid. Derived from fuel mix, ACI is usually only available from unvalidated 3rd parties (e.g. Electricity Maps [64]) in a fraction of the world's grids.
- **Grid price or Price (\$/MWh)** is power price in a grid or region (e.g. "hub price"). Renewable generator often bids low, causing power price to be correlated with carbon intensity (Figure 2). Price information is widely available in day-ahead and real-time markets.
- **Locational marginal price or LMPrice (\$/MWh)** is the price at a specific node in the power grid. LMPrice reflects local properties such as nearby renewables and grid transmission constraints. LMPrice is widely-available.

Recently, researchers and companies [59, 90] have proposed marginal carbon intensity as a metric. We do not consider it, as it is not widely available, and tied to proprietary market strategy.

5.2 Modeling Datacenter Capacity Flexibility

Datacenter capacity flexbility defines the variation structure of capacity. We assume that every datacenter can adjust capacity (**cap**) within a **dynamic range** and defer workload (**backlog**), but must catch up within a 24-hour day (Figure 5), formally:

$$cap_{min} \leq cap_{i,t} \leq cap_{max},\ \forall i, t \tag{1}$$

$$backlog_{i,t} = backlog_{i,t-1} + (avgCap - cap_{i,t}) \tag{2}$$

$$backlog_{i,24} = 0,\ \forall i \tag{3}$$

These flexibility constraints are common in datacenter carbon-aware capacity adaptation studies [4, 56, 78].

Datacenter capacity variation can harm a datacenter's *computation performance* [95] and even power markets. We consider a **step size** limit that bounds hour-to-hour datacenter capacity change:

$$|cap_{i,t} - cap_{i,t-1}| \leq stepSize,\ \forall i, t \tag{4}$$

DC attributes used are shown in Table 1. The resource utilization and capacity assumptions are typical of hyperscale datacenters

[1] Sometimes, these are called grid "signals" for load adaptation.

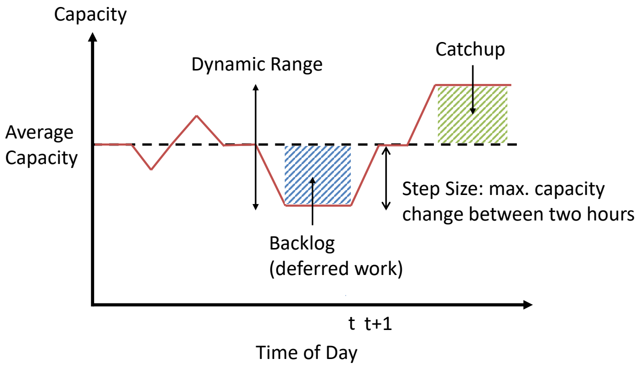


Figure 5: Datacenter Capacity Flexibility Model.

[3, 34, 69, 85]. We present results from the [0.4, 1.0] dynamic range which combines the largest potential benefits, but also overshifting challenges. High fractions of deferrable workload reflect the industry's published workload papers. Google's BorgTNG trace [85] shows flexible jobs with 24-hour completion SLO (service level objective) make up about 40% of the resource usage; delay-insensitive VMs (virtual machines) account for about 68% of resource usage among Microsoft Azure VM workload [21]; in Meta, 60% of batch jobs can be flexibly scheduled within a day [4]. It's likely that the workload flexibility continues to grow due to emerging batch workloads like machine learning model training. In addition, the potential deployment of long-duration energy storage is a complementary solution to datacenter capacity adaptation [49].

Table 1: Configurations of Datacenter Attributes

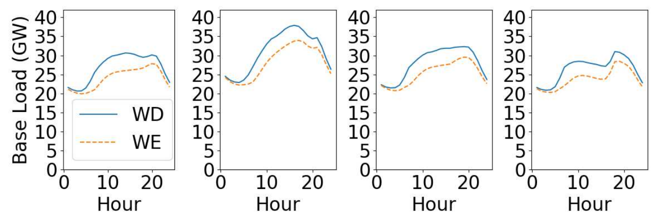
Attribute	Configuration(s)
Maximum Capacity	200 MW
Average Utilization Level	70%
Average Capacity (*avgCap*)	140 MW
Dynamic Range	[0.6, 0.8], [0.4, 1.0]
Step Size	10, 20, 40, 80, 120 MW/h

5.3 Power Grid Model

We evaluate DC adaptation approaches coupled to a realistic grid model. Grid operation (day-ahead planning or real-time operation) is simulated by solving the direct-current optimal power flow (DC-OPF) problem in [50, 56] and Appendix A, which minimizes the grid dispatch cost in one-day time horizon with hourly intervals, subject to typical grid constraints. The grid metrics for carbon optimization (ACI, Price, LMPrice) are derived from the OPF solutions. With lower generation costs and curtailment penalties that encourage use, renewable generators produce low prices when dispatched at the margin, capturing the the correlation between carbon metrics and power price in the real world (Figure 2).

The grid topology is a reduced California power system (CAISO) consisting of 225 buses, 375 transmission lines, 130 thermal generators (31.2 GW total capacity), 11 non-wind renewable power plants, 5 wind power plants, and 40 loads. Power can also be imported at 5 boundary buses. This model is originally from [72] and has been used to assess the impact of dynamic datacenter capacity management in [50, 56]. We select wind as the major renewable generation source as it presents more intra-day variation that leads to more diverse capacity adaptation behaviors. Besides, it can model the wind-dominant power grids such as ERCOT (Electric Reliability Council of Texas) and SPP (Southwest Power Pool). To get higher wind penetration, the wind generation are scaled up equally at current sites, assuming those sites can be expanded or equipped with higher-capacity wind turbines [76].

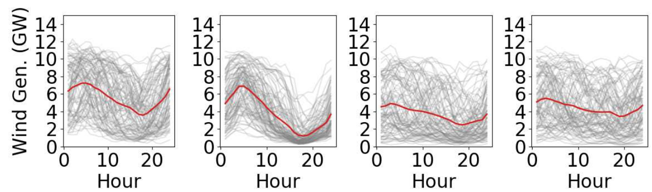
There are 8 base load, imports, and non-wind renewable generation profiles that cover the four seasons (Spring, Summer, Fall, Winter) and weekday/weekend (WD/WE). Figure 6 shows how each load profile varies in a day, with the average spanning from 23,780 MW (WinterWE) to 31,089 MW (SummerWD). We assume high accuracy of base load and renewable generation forecast, so the day-ahead and real-time OPF share the same deterministic load and renewable generation profiles.



(a) Spring (b) Summer (c) Fall (d) Winter

Figure 6: Grid Base (Non-DC) Load Profiles.

To reflect the impact of wind variation, for each season, we use 100 wind scenarios (Figure 7) shared by the weekday and weekend. The wind generation tends to be higher in the late night and early morning, which is a misalignment with load and can be opportunities for datacenter capacity adaptation.



(a) Spring (b) Summer (c) Fall (d) Winter

Figure 7: Used Wind Scenarios (15% Penetration). Red lines represent average wind scenarios for each season.

Datacenters are added to random buses in the grid (with loads added to the base load profiles), which reflects the fact that datacenter site selection is more based on business considerations external to the power grid (e.g. tax breaks, jobs, internet hookups, etc.).

5.4 Evaluation Metrics

Our evaluation metrics cover both datacenter goals and impacts on the grid and other grid customers:

- **Datacenter Carbon Reduction.** We vary only datacenter capacity, and thus attribute the grid carbon reduction to DC capacity adaptation, reporting it as percentage reduction in datacenter operational carbon emissions:

$$\frac{(gridCarbon_{fixed-cap} - gridCarbon_{adaptation})}{datacenterCarbon_{fixed-cap}} * 100\%$$

  and grid carbon emissions are calculated as:

$$gridCarbon = \sum_{t=1}^{24} gen_{f,t} * emissionRate_f$$

  where $gen_{f,t}$ is generation from fuel f in the t-th hour. Fuel emission rates are from US EPA eGrid database [6, 42] and listed in Appendix B. *datacenterCarbon* with fixed-capacity DCs is calculated using the grid emission rate.
- **Grid Dispatch Cost ($)** is the objective for DC-OPF minimization and a figure of merit for grid operation.
- **Datacenter Average Power Price ($/MWh)** is how much the datacenter would pay for power considering the location of datacenter and the time when power is consumed. For multiple datacenters, we compute the average.
- **Non-datacenter Customer Average Power Price ($/MWh)** is the average power price across grid customers other than datacenters, weighted by power demand.
- **Datacenter Average Capacity Variation (MW/h)** is defined as the average change in power capacity of a datacenter between adjacent one-hour periods, the lower the better. More formally:

$$\frac{1}{23}\sum_{t=2}^{24} |cap_{i,t} - cap_{i,t-1}|$$

  where $cap_{i,t}$ denotes the capacity of datacenter i at time t.

Conventional datacenters operate at fixed capacity; we use this as the baseline when examining the impacts of datacenter capacity adaptation.

5.5 Experiment Setup

Given the 100 wind scenarios for each day type, our simulation is equivalent to simulating a total of 800 days [72]. For each day, we vary the wind penetration level and simulate datacenter operation using different capacity adaptation algorithms. The results reported in Section 6 are the average of 8 day types (weighted for the number of weekdays and weekend days) and varied wind scenarios.

We used Julia 1.5.2 with JuMP v0.21.5 [26] to implement the grid simulation and solved grid OPF with Gurobi Optimizer v9.0 [38].

6 EVALUATING DATACENTER CAPACITY ADAPTATION APPROACHES

We explore datacenter capacity adaptation algorithms for each of the three approaches in Section 4. Evaluation uses a full grid simulation and results are compared to datacenters with fixed-capacity (no capacity adaptation). While we have studied many scenarios, for clarity we show a single representative scenario (30 200-MW datacenters and dynamic range of [0.4, 1.0]). In this scenario, datacenters are 10% of grid load, less than in leading-edge grids, but realistic for dozens of grids throughout the world in the near future.

6.1 No Coordination (Local Scope)

With local, online capacity adaptation, each datacenter makes real-time capacity decisions independently based on current and future grid metrics. We employ a dynamic programming algorithm that makes hourly decisions using current value of metric and its daily average, and refer to local adaptation based on different metrics by "**<Metric> (Avg)**". The algorithm selects amongst $\{avgCap, avgCap \pm \text{dynamic range}/2\}$ to minimize the expectation:

$$cap_{i,t} * metric_{i,t} + (backlog_{i,t-1} + avgCap - cap_{i,t}) * metric_i$$

which can be either carbon emissions or power cost. The *backlog* is then updated given the determined power capacity.

Datacenter capacity adaptation is coupled to grid dispatch (OPF optimization), as below, where datacenters are denoted by i, and hours by t:

1. For all i,t: $cap_{i,t} = avgCap$ (neutral initial condition).
2. Solve grid OPF with $\{cap_{i,t}\}$, defining $metric_{i,t}$ as day-ahead information.
3. At the beginning of $t = 1, ..., 24$-th hour, each datacenter adapts capacity $cap_{i,t}$ based on the metrics.
4. Then the grid solves OPF with updated datacenter capacities $\{cap_{i,t}\}$ (real-time operation).
5. This new OPF solution redefines $metric_{i,t}$ (realized) and for the future (i.e. $[t + 1, ..., 24]$).

6.1.1 Comparing Grid Metrics. We evaluate the effectiveness of different grid metrics used for adaptation—average carbon intensity (ACI), grid price (Price), and locational marginal price (LMPrice). LMPrice consistently outperforms ACI and Price (Figure 8).

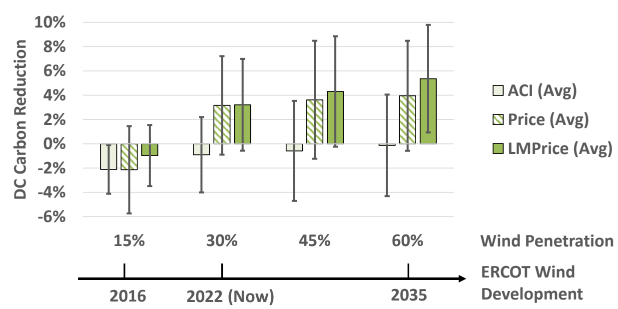


Figure 8: DC Carbon Reduction with Adaptation based on Carbon Intensity, Grid Price, and Locational-Marginal Price.

The x-axis in Figure 8 is (15%–60% wind penetration) and represents expected progress of renewable generation; for calibration we use the expected timing of this change for ERCOT as a reference. We use this x-axis in many figures. We also plot standard deviation across daily variation in renewable generation (wind scenarios) as "whiskers". Each specific wind scenario typically produces correlated results for different algorithms, so these whiskers capture variability, not uncertainty.

At 15% wind, online adaptation fails to reduce carbon-emissions because generation supply is tight, and oversubscription happens during low-carbon periods. Higher wind penetration provides more opportunity, enabling online adaptation using LMPrice to reduce datacenter carbon emissions by 5.4% (60% wind), consistently outperforming ACI (+5.4%) and Price (+1.4%), Our broader studies show that price metrics work better generally, and finer-grained pricing works best.

To illustrate, Figure 9 shows a single-day timeline. The graphs show that the grid-wide signals, ACI and Price, create uniform, lockstep datacenter behavior, maximizing the load swings for the grid. In contrast, the locational metric, LMPrice, produces diverse behaviors as each datacenter reacts to local pricing that reflects grid constraints such as transmission and generator ramping.

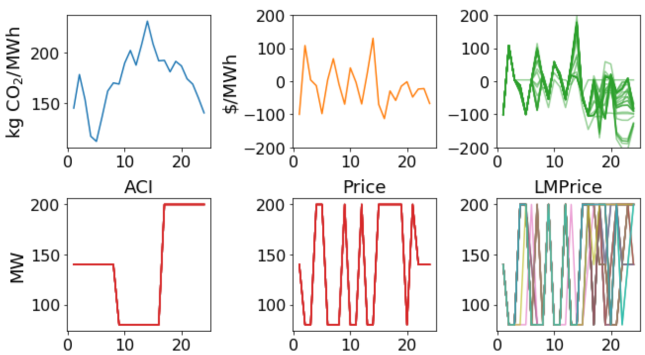


Figure 9: Global vs. Local metrics (LMPrice) create different DC capacity adaptation (bottom). Local produces DC adaptation variety. (Fall Weekday, 30% Wind Penetration)

6.1.2 Improving Local Online Adaptation. Using LMPrice, we next explore how to improve local-scope adaptation by adding finer-resolution price , better forecasts, and smoothing capacity changes.

Using the day-ahead hourly locational prices at the DC site (price array $\{\text{predLMP}_{i,t}\}, t = 1, ..., 24$), in the j-th hour, datacenter i performs dynamic programming on the price array $\{p_{i,t}\}$ with:

$$p_{i,t} = \begin{cases} \text{LMP}_{i,j}, & \text{if } t = j \text{ and } j \neq 1 \\ \text{predLMP}_{i,t}, & \text{otherwise} \end{cases}, \ t = j, ..., 24$$

where $\text{LMP}_{i,j}$ is the real-time price after $\{cap_{i,j-1}\}$ are set. The dynamic programming algorithm produces a capacity array based on the following recurrence formula:

$$cost_i(n, t, cap) = cap * p_{i,t} + \min_{cap'}\{cost_i(n + cap - avgCap, t - 1, cap') \mid |cap - cap'| \leq stepSize\} \quad (5)$$

where $cost_i(n, t, cap)$ denotes the minimum power cost of the subproblem ending at t-th hour with backlog n and capacity level cap in the t-th hour. Following the convention of dynamic programming, the resulting capacity array is obtained through backtracking from $min(\{cost_i(0, 24, cap) \mid cap_{min} \leq cap \leq cap_{max}\})$, which is the optimal cost subject to the constraints defined in Section 5.2. The datacenter takes the first element of the solution array (the last in backtracking) as $cap_{i,t}$—similar to the receding horizon control [51] but with a fixed horizon. This algorithm models:

Forecasts (Future Information). We use hourly LMPrice from the day-ahead market as a price forecast. It is refined in other markets (e.g. hourly, 15-minute, 5-minute, real-time) as the time approaches until the final OPF determines the dispatch, prices, etc. Adding this information—the full 24 hours of day-ahead prices—is reflected in **LMPrice (Hourly)**.

Step Size (Smoothed Capacity). Online adaptation can cause large datacenter capacity fluctuations, harming both datacenter computing efficiency and the grid (generation dispatch, carbon intensity, price). To smooth these large capacity changes, we limit hour-to-hour DC capacity change with a maximum step size.

Using detailed future price information, we empirically determined that a step size of 40 MW/h for LMPrice (Hourly), and unbounded for LMPrice (Avg) that yield the largest carbon reduction. 50% of the improvement from LMPrice (Avg) to LMPrice (Hourly) is due to price information and 50% to step size (Figure 10).

In Figure 10, **LMPrice (6hr + Avg)** reflects progressively introduced future information, which is given hourly information for the next 6 hours but only daily average after that. For 30% wind penetration, 6 hours' information enables 2.6% reduction and 24 hours' a 4.2% reduction in datacenter carbon emissions. These grow to 3.5% and 4.6% respectively at 60% wind penetration. Ultimately LMPrice (Hourly) delivers 10% of DC carbon reduction.

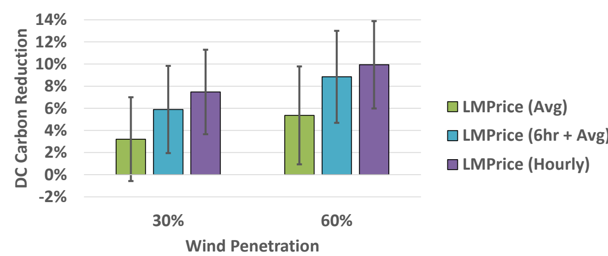


Figure 10: Datacenter Carbon Reduction with varied future LMPrice information (Day average, 6hr detailed, 24hr detailed).

6.2 Coordinating a Group of Datacenters

Independently controlled DCs can react together, when decisions are based on grid-wide or other strongly-correlated metrics, producing a large aggregate power capacity change. We have seen LMPrice is better, but even its correlation across sites can produces synchronized DC capacity changes that are difficult for the grid to manage. Addressing this, we add an external limiter, called "coordinator", to mitigate the harm. In **LMPrice (Avg)-Coord**, each datacenter runs independent online control algorithm and then submits their adaptation to a coordinator. The coordinator limits total power change for a set of datacenters, using a quota:

(1) Generate a random permutation of datacenters.
(2) For each datacenter, if $change \leq quota$ then accept, $quota = quota - change$; else, reject.

If local controllers were unable to get their requested change, they must update their backlogs accordingly.

Figure 11 shows DC carbon reduction with varied coordinator quotas (3600–1200 MW/h). Each line represents a different wind penetration level. Coordination improves performance, mitigating

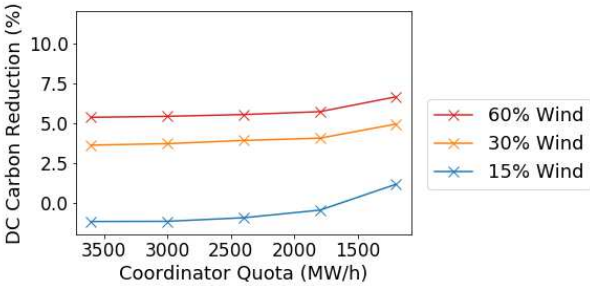


Figure 11: Datacenter Carbon Reduction vs. Coordinator Quota for three wind penetration levels.

overshifting harm at 15-60% wind penetration: on each line, all the points reflect greater carbon reduction than the leftmost one—LMPrice (Avg), and steady improvement as the quota is tightened. The benefit for datacenters is larger when the overshifting is more evident under tight generation (15% wind penetration), improving DC carbon reduction by 2.3%. These benefits are smaller than those achieved by exploiting future price information, such as LMPrice (Hourly), which yields a 50% higher DC carbon reduction.

Multiple Coordinators. In many geographic areas, there are multiple cloud providers (e.g. Northern Virginia, Texas, Ireland, or China's Ningxia), and each cloud provider has multiple datacenter sites in that area. As competitors, they may not be willing to share a coordinator. To model this multi-coordinator scenario, we assume there are 2 or 3 coordinators in the grid, each coordinating 10 or 15 DCs. We use an overall quota of 1200 MW/h, and divide it equally across the coordinators.

Increasing the number of coordinators further decreases DC carbon emissions. The reason for this is narrower coordinator scope with datacenter load quantization increases the smoothness of total datacenter load, but the improvements are small. 3 coordinators produce improvements up to 0.7% of DC carbon emissions, achieving 7.3% of DC carbon reduction at 60% wind penetration.

6.3 Coordination with the Grid: Capacity Plan Sharing

Adaptive datacenters cause grid problems as their large power changes are unpredictable and strain generator ramp constraints. Further, unplanned adaptation can produce rapid changes in compute capacity, making it difficult for cloud resource managers to be efficient. In view of these insights, we expand the space scope to datacenter-grid cooperation and time scope to day-ahead (planning), proposing a new approach—**PlanShare**: datacenters create a 24-hour adapted capacity plan based on LMPrice in day-ahead grid market [44] ahead of operating day, and then share the plan with the grid. This allows the grid to optimize globally based on the DC information. Formally, with datacenters denoted by i and hours denoted by t:

(1) For all i,t: $cap_{i,t} = avgCap$ (neutral initial condition).
(2) Solve grid OPF with $\{cap_{i,t}\}$, defining initial $\text{LMP}_{i,t}$ as day-ahead information.
(3) Each DC makes 24-hour adaptation plan using LMPrice (Hourly)'s dynamic programming algorithm and shares it with the grid.
(4) Solve grid OPF for $[1, ..., 24]$ with adapted $\{cap_{i,t}\}$ to model the next day's operation—the datacenter must follow the full-day capacity plan it shares with the power grid.

Figure 12 compares PlanShare and the local, online approaches. The results show the benefits of plan sharing with the grid. Using essentially the same adaptation algorithm (comparison of behaviors shown in Figure 13), and on less accurate information, by working with the grid, PlanShare reduces DC carbon emissions by up to 12.6% (1.6% grid carbon reduction if normalized to grid carbon emissions). This is 1.26x better than the best local, online adaptation result, and the advantage is even higher (1.56x) with today's wind penetration levels (30%). By contributing its adaptation plan to the grid optimization—in advance and as a committed schedule—PlanShare dramatically improves the datacenter carbon emissions reduction that can be achieved.

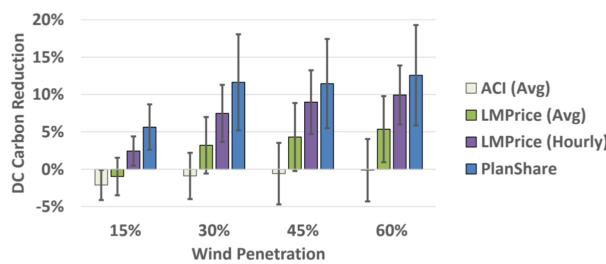


Figure 12: DC Carbon Reduction for four capacity adaptation schemes. PlanShare outperforms all Online approaches and under all wind penetration levels.

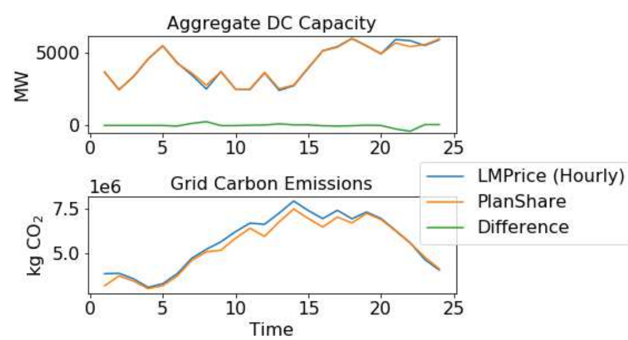


Figure 13: PlanShare produces similar aggregate DC capacity (Top) but sharing information with the grid produces 8x larger DC emissions reduction (Bottom). (Fall WD Example)

Sensitivity to Length of Shared Plan. Having demonstrated the benefits of a practical scheme (24-hour day-ahead plans are available in many power grids), we explore **how much DC plan information is needed by the grid?** We vary the length of the DC capacity plan shared from 1 to 24 hours, reporting results in Figure 14. Interestingly, a single hour advance plan is enough for PlanShare to match the performance of LMPrice (Hourly), the best online approach, despite the fact that PlanShare has no online adaptation; the full 24-hour capacity plan is fixed. As the length of plan is increased to 3, 6, and 12 hours, the benefit of plan sharing increases significantly, reaching the large benefits previously highlighted in Figure 12.

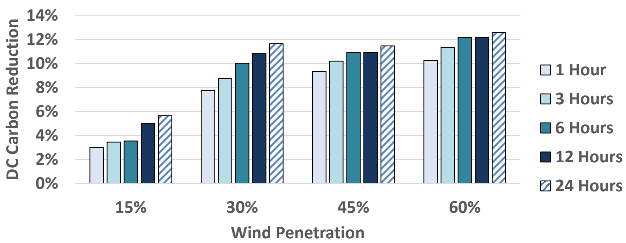


Figure 14: DC Carbon Reduction vs. Varying Plan Length: PlanShare with 1-24hr plan.

6.4 Datacenter Adaptation Impacts beyond Carbon

Datacenter capacity adaption produces other impacts on the grid customers and datacenter operation. We study several here.

Grid Dispatch Cost. We compare the grid dispatch cost of different approaches in Figure 15. Local, online datacenter adaptation approaches can increase grid dispatch cost at lower wind penetration (up to 6% increase by ACI (Avg)). PlanShare successfully eliminates this grid performance damage, decreasing grid dispatch cost by as much as 2.5%. The reduction is attributed to decreased generation cost and renewable curtailment penalties. Datacenter capacity adaptation and grid-wide optimization given the shared plans enable this.

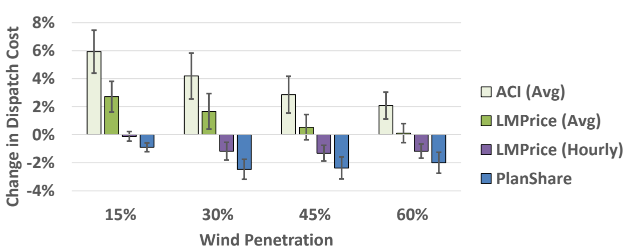
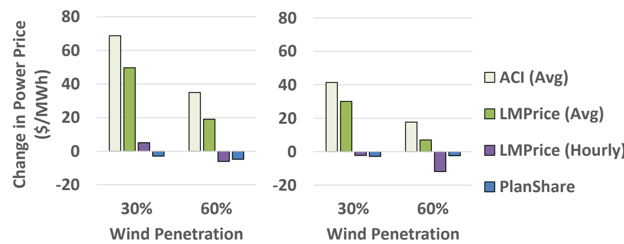


Figure 15: Grid Dispatch Cost vs. Adaptation Scheme—PlanShare reduces grid dispatch cost most.

Datacenter Power Cost. Adaptive datacenters affect power pricing in the grid. In Figure 16a, results show that local, online adaptation can cause significant power prices increases of \$20 to \$50/MWh, corresponding to 59%–490% increase in power cost. This is because locally controlled adaptation clashes with grid constraints and dynamics (overshifting in Section 3), and it is likely a major deterrent for datacenter adoption. In contrast, sharing the datacenter's capacity plan in advance as in PlanShare decreases average power price stably, up to 30% compared with the fixed-capacity scenario.



(a) Datacenters. **(b) Other (Non-DC) Customers.**

Figure 16: Change in Average Power Price for Different Grid Customers.

Non-DC Customer Power Cost. In Figure 16b, We explore how adapting datacenter capacity impacts other (non-DC) customers. Local approaches, including ACI (Avg) and LMPrice (Avg), significantly increase the price for non-DC customers, especially for lower wind penetration with less excess renewable generation (LMPrice (Hourly) also increases price at 15% wind). Datacenters as growing consumers of power are already subject to growing scrutiny and negative publicity. Pricing hard to other customers has the potential to cause a backlash, so datacenters should be careful about deploying such local, online adaptive capacity control. PlanShare avoids this price harm for non-DC customers at all wind penetration levels.

Datacenter Capacity Variation. Capacity variation is a critical concern for datacenter operators as it affects workload efficiency of the available capacity. Using the average capacity variation metric (see Section 5), Figure 17 shows DC carbon reduction on y-axis and capacity variation on x-axis. As before, DC carbon reduction is relative to the fixed-capacity scenario, and capacity variation (MW/h) is normalized to the maximum capacity (200 MW). An ideal adaptation approach would fall in the upper-left corner (high carbon reduction, low capacity variation). Each line connnects results as wind penetration increases for a given adaptation approach.

The plot clearly shows how the higher-performing adaptation techniques exploit increased changes to adapt to the changing grid carbon properties. The progression from ACI (Avg) to LMPrice (Avg) to LMPrice (Hourly) shows a tradeoff of carbon reduction for online capacity variation. Achieving the greatest carbon reduction, PlanShare does require greater capacity variation. However, it's worth noting that PlanShare fixes the capacity plan in advance, so the resource manager has a statically known resource schedule at the start of the day, facilitating compute workload scheduling. Scheduling studies and other proposals [78, 94, 95] argue for the benefits of known capacity information. For example, with 24-hour capacity information, the cloud workloads see little performance (e.g. goodput, job wait time) degradation at 60% dynamic range compared with the fixed capacity scenario[94].

Table 2: Summary of Datacenter Capacity Adaptation Approaches

	ACI (Avg)	LMPrice (Avg)	LMPrice (Hourly)	PlanShare (1 hour)	PlanShare (24 hours)
DC Carbon Reduction	–	neutral	+	+	++
DC Power Cost	– –	–	neutral	+	++
DC Capacity Variation	+	+	neutral	neutral	neutral
Grid Dispatch Cost	–	neutral	+	+	++
Other Customers (non-DC) Power Cost	– –	–	neutral	+	++

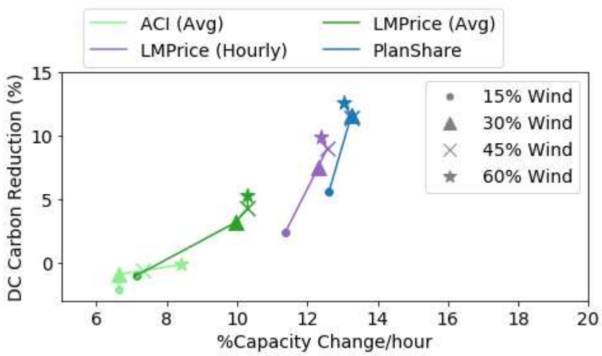


Figure 17: DC Carbon Reduction and Capacity Variation.

6.5 Summary

Table 2 summarizes the impacts of datacenter capacity adaptation in different metrics, where "+" means advantage and "–" means disadvantage. The weakest performance plan sharing approach—PlanShare with 1 hour's shared plan—matches and outperforms all of the other approaches. And, as we can see in Figure 14, PlanShare with 24-hour adaptation plan outperforms others significantly and is by far the best, delivering the greatest DC carbon reduction while enabling datacenter resource managers to have known capacity plans. To conclude, PlanShare satisfies both datacenter (carbon and cost reduction, computing efficiency) and power grid objectives (cost reduction and fairness).

7 DISCUSSION AND RELATED WORK

To the best of our knowledge, this is the first paper that explores the coordination scope for datacenter capacity adaptation and proposes sharing datacenter capacity plans with the grid to reduce operational carbon emissions. Given LMPrice metric availability, and cloud practice making day-ahead adapted capacity plan [78], our PlanShare approach is feasible. Some cloud providers may be reluctant to share their daily capacity plan with the grid for proprietary reasons. But it's worth pointing out that each datacenter's local grid (e.g. Dominion Energy, PGE) already knows DC's historical power consumption, going back days, months, and years. Further, if the cloud DCs wanted to intentionally mask their compute load, they could still do so with on-site batteries or even generators.

We review related work below:

Datacenter Capacity Adaptation (Shaping). Early ideas like "follow-the-moon""chase-the-wind" propose to shift datacenter workload to a time or place with low energy prices or carbon emissions [1], exploiting variations of grid dynamics. The DC capacity adaptation explored in this paper addresses a widely studied subset of these ideas—temporal load shaping or shifting, which typically employ sophisticated online control or optimization techniques [24, 57, 60, 63]. An additional variant is to manage colocated energy storage [7, 32, 81, 87, 93]. This work assumes datacenters are small loads (grid trace-based studies or DC-only evaluation), and doesn't model the impact of DC dynamic capacity changes on the grid, not to mention coupled impacts on carbon, prices, generation, etc. In this paper, a collection of hyperscale datacenters (200 MW each) coupled to a power grid model is essential to capturing their direct impacts on grid dispatch.

Recent cloud industry goals include "24×7" or "100/100/0" hourly matching of DC power consumption with carbon-free generation [35, 46]. These efforts explore application and resource-management load shifting based on varying power carbon-intensity[78], do not assess impact on power grids. For example, Google's *carbon-aware computing* [78], creates day-ahead capacity plan (called a *virtual capacity curves* (VCC)) to enable efficient DC resource management. But the VCC is not shared with the power grid.

Another dimension of load shifting is spatial or geographic, with similar goals [37, 52, 61, 77, 98, 99, 101, 102]. This promising direction is beyond the scope of our paper.

Datacenters in Demand Response. Datacenters' ability to delay workload (e.g. defer batch jobs) can be used to participate in demand-response programs, reducing capacity during peak periods according to requests from the grid. Such participation can reduce DC power cost, and research has explored how to balance this benefit while respecting service-level objectives (SLO) [53, 62]. Further efforts in this area design sophisticated markets that incentivize DC operators and even their colocation tenants to participate in demand response [14, 15, 84, 96, 100].

Demand response is designed for emergency reduction in load to protect grid stability. As a result, actions are rare, and the power reduction is small relative to the total power consumption [67]. In contrast, we consider multiple datacenters' active, continuous capacity adaptation for reducing operational carbon emissions with dynamic range up to 60% of capacity.

Grid-coupled Datacenter Adaptation. As early as 2015, Yang, Wolski, and Chien proposed *Zero-carbon Cloud*, a novel approach that proposed building datacenters as dispatchable loads controlled by the grid to harness excess carbon-free power [17, 91, 92], and the grid benefits including decreased dispatch cost and renewable curtailment are reported in a study including a grid model [50]. Commercial efforts at gigawatt-scale based on Zero-carbon Cloud ideas are under construction in ERCOT/Texas [2]. Other efforts include exploring the grid benefits of temporal or spatial load shifting [59, 68, 97]. [58] studies what grid metrics can effectively guide carbon-aware spatial load shifting. While also showing locational marginal price is effective, they claim locational mariginal carbon emissions is better, which, however, is not broadly available today.

This paper builds on the insights of [56]. The authors showed that without modeling the grid's dynamics, the carbon emission projections could be significantly wrong. Thus, it's necessary to include grid models in studies of large-scale temporal workload shifting. However, that work provides no solution to coordinated management of datacenters and power grids.

8 SUMMARY AND FUTURE WORK

Cloud providers are adapting datacenter capacity to renewable generation to reduce carbon emissions. However, for today's large cloud datacenters, the numerous prior techniques based on independent, online control fail to reduce emissions and can harm the grid. To find a robust solution, we explore the coordination of adaptation, varying scopes in time and space. With the local coordination scope, locational marginal price (LMPrice) is identified as a widely available and the most effective grid metric for datacenter capacity adaptation. Expanding the scope to grid-wide coordination and day-ahead planning, we propose a solution—PlanShare, where each datacenter creates a capacity plan based on day-ahead grid metric, and then shares it with the grid. This approach enables DCs to achieve greater emissions reduction (12.6%), lower average power prices (-30%), and more predictable capacity from planning (and thus better internal utilization). PlanShare even eliminates harm to other customers in the grid. We are optimistic that as power grids are increasingly dominated by renewable generation the techniques studied here will enable datacenters to balance the efficient delivery of cloud computing with helping balance and decarbonize the power grids. As datacenters grow—fueled by artificial intelligence, pervasive intelligent control, commerce, and entertainment—beyond 10 and 20 percent of power grid load, these techniques will be essential not only for the power grid, but for the continued growth of computing.

Several exciting future directions for research include: First, how to create compute load flexibility and respond to capacity change, while respecting service-level objectives (SLO)? The resource managers and applications today lack a clear cost metric, and further it's unclear what types and extents of flexibility are possible or valuable. Second, we have focused on temporal capacity adaptation, but it's also interesting to explore spatial shifting combined with day-ahead capacity plan schedules. Third, how can we balance datacenter privacy considerations with the clear benefit of sharing capacity plan information? Finally, as the world progresses to higher levels of renewable generation and computing power consumption, it will be essential to reconsider the techniques proposed here.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for the insightful reviews, including who reviewed the earlier versions of this paper. This work is supported in part by NSF Grants CMMI-1832230, OAC-2019506, and the VMware University Research Fund. We also thank the Large-scale Sustainable Systems Group members for their support of this work!

REFERENCES

[1] 2009. Follow the Sun, Wind, Moon. https://www.vertatique.com/cloud-computing-starting-follow-sunwindmoon.

[2] 2018. Lancium. https://www.lancium.com. A startup company, building zero-carbon cloud computing resources..

[3] 2019. Clicking Clean Virginia: The Dirty Energy Powering Data Center Alley. https://www.greenpeace.org/usa/reports/click-clean-virginia/.

[4] Bilge Acun, Benjamin Lee, Kiwan Maeng, Manoj Chakkaravarthy, Udit Gupta, David Brooks, and Carole-Jean Wu. 2022. A Holistic Approach for Designing Carbon Aware Datacenters. *arXiv preprint arXiv:2201.10036* (2022).

[5] U.S. Energy Information Administration. 2022. Electric Power Monthly. https://www.eia.gov/electricity/monthly/epm_table_grapher.php?t=epmt_5_03

[6] United States Environmental Protection Agency. 2020. Emissions & Generation Resource Integrated Database. https://www.epa.gov/egrid/data-explorer

[7] Sohaib Ahmad, Arielle Rosenthal, Mohammad H. Hajiesmaili, and Ramesh K. Sitaraman. 2019. Learning from Optimal: Energy Procurement Strategies for Data Centers. In *Proceedings of the Tenth ACM International Conference on Future Energy Systems* (Phoenix, AZ, USA) *(e-Energy '19)*. Association for Computing Machinery, New York, NY, USA, 326–330. https://doi.org/10.1145/3307772.3328308

[8] Amazon. 2022. 2021 Environmental Report. Available from https://sustainability.aboutamazon.com/reporting-and-downloads.

[9] Amazon Web Services (AWS). [n. d.]. Amazon Global Datacenters.

[10] Noman Bashir, Tian Guo, Mohammad Hajiesmaili, David Irwin, Prashant Shenoy, Ramesh Sitaraman, Abel Souza, and Adam Wierman. 2021. Enabling Sustainable Clouds: The Case for Virtualizing the Energy System. In *Proceedings of the ACM Symposium on Cloud Computing*. 350–358.

[11] Lori Bird, M Milligan, and Debra Lew. 2013. *Integrating Variable Renewable Energy: Challenges and Solutions*. Technical Report. NREL.

[12] Bloomberg. 2018. European Union Aims to Be First Carbon Neutral Major Economy by 2050. *Fortune* (November 2018).

[13] John Campbell. 2022. Data centres used 14% of Republic of Ireland's electricity use. https://www.bbc.com/news/world-europe-61308747

[14] Niangjun Chen, Xiaoqi Ren, Shaolei Ren, and Adam Wierman. 2015. Greening multi-tenant data center demand response. *Performance Evaluation* 91 (2015), 229–254.

[15] Shutong Chen, Lei Jiao, Lin Wang, and Fangming Liu. 2019. An online market mechanism for edge emergency demand response via cloudlet control. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*. IEEE, 2566–2574.

[16] Andrew A Chien. 2020. Characterizing Opportunity Power in the California Independent System Operator (CAISO) in Years 2015-2017. *Energy and Earth Science* 3, 2 (December 2020). Also available as University of Chicago, Computer Science TR-2018-07, https://newtraell.cs.uchicago.edu/research/publications/techreports.

[17] Andrew A Chien, Richard Wolski, and Fan Yang. 2015. The Zero-Carbon Cloud: High-Value, Dispatchable Demand for Renewable Power Generators. *The Electricity Journal* (2015), 110–118.

[18] Andrew A. Chien, Fan Yang, and Chaojie Zhang. 2018. Characterizing Curtailed and Uneconomic Renewable Power in the Mid-continent Independent System Operator. *AIMS Energy* 6, 2 (December 2018), 376–401.

[19] Microsoft Corporation. 2021. Microsft 2021 Environmental Sustainability Report. Available from https://www.microsoft.com/en-us/corporate-responsibility/sustainability/report.

[20] Olivier Corradi. 2022. Marginal vs average: which one to use in practice? https://www.electricitymaps.com/blog/marginal-vs-average-real-time-decision-making

[21] Eli Cortez, Anand Bonde, Alexandre Muzio, Mark Russinovich, Marcus Fontoura, and Ricardo Bianchini. 2017. Resource Central: Understanding and Predicting Workloads for Improved Resource Management in Large Cloud Platforms. In *Proceedings of the 26th Symposium on Operating Systems Principles (SOSP '17)*. 153–167. https://doi.org/10.1145/3132747.3132772

[22] CPUC. [n. d.]. California Public Utilities Commission (CPUC). https://www.cpuc.ca.gov/rps.

[23] Wei Deng, Fangming Liu, Hai Jin, Bo Li, and Dan Li. 2014. Harnessing renewable energy in cloud datacenters: opportunities and challenges. *iEEE Network* 28, 1 (2014), 48–55.

[24] Hui Dou, Yong Qi, Wei Wei, and Houbing Song. 2017. Carbon-aware electricity cost minimization for sustainable data centers. *IEEE Transactions on Sustainable Computing* 2, 2 (2017), 211–223.

[25] William Driscoll. 2022. California law would target 90% renewable and zero-carbon electricity by 2035. https://pv-magazine-usa.com/2022/09/06/california-law-would-target-90-renewable-and-zero-carbon-electricity-by-2035/

[26] Iain Dunning, Joey Huchette, and Miles Lubin. 2017. JuMP: A modeling language for mathematical optimization. *SIAM review* 59, 2 (2017), 295–320.

[27] Dominion Energy. 2018. 2018 Virginia Integrated Resource Plan. https://rga.lis.virginia.gov/Published/2018/RD249

[28] Dominion Energy. 2020. 2020 Virginia Integrated Resource Plan. https://www.dominionenergy.com/-/media/pdfs/global/2020-va-integrated-resource-plan.pdf

[29] Dominion Energy. 2021. 2021 Update to the 2020 Integrated Resource Plan. https://www.dominionenergy.com/-/media/pdfs/global/company/2021-de-integrated-resource-plan.pdf

[30] Mary Jo Foley. 2022. Cloud revenues power Microsoft's $51.7 billion Q2 in fiscal year 2022. https://www.zdnet.com/article/microsoft-cloud-revenues-power-microsofts-51-7-billion-second-fy22-quarter/.

[31] Robbie Galvin. 2021. Data Centers Are Pushing Ireland's Electric Grid to the Brink. https://gizmodo.com/data-centers-are-pushing-ireland-s-electric-grid-to-the-1848282390

[32] Íñigo Goiri, William Katsak, Kien Le, Thu D Nguyen, and Ricardo Bianchini. 2013. Parasol and greenswitch: Managing datacenters powered by renewable energy. In *ACM SIGARCH Computer Architecture News*. ACM, 51–64.

[33] Caroline Golin and Nick Pearson. 2022. A policy roadmap for 24/7 carbon-free energy. https://cloud.google.com/blog/topics/sustainability/a-policy-roadmap-for-achieving-247-carbon-free-energy

[34] Google. [n. d.]. About Google Datacenters. https://www.google.com/about/datacenters/.

[35] Google. 2018. *Moving toward 24x7 Carbon-Free Energy at Google Data Centers: Progress and Insights.* Technical Report. Google.

[36] Google. 2022. Google 2021 Environmental Report. Available from https://www.gstatic.com/gumdrop/sustainability/google-2022-environmental-report.pdf.

[37] V. Gupta, P. Shenoy, and R. K. Sitaraman. 2018. Efficient solar provisioning for net-zero Internet-scale distributed networks. In *2018 10th International Conference on Communication Systems Networks (COMSNETS)*. 372–379.

[38] LLC Gurobi Optimization. 2020. Gurobi Optimizer Reference Manual. http://www.gurobi.com

[39] GWEC. 2016. *Global Wind Report: Annual Market Update.* Technical Report. Global Wind Energy Council. Documents curtailment around the world.

[40] Siqi Han. 2015. The Wind is Wasted in China. https://www.wilsoncenter.org/.

[41] Kevin Imboden. 2021. 2022 Global Data Center Market Comparison. https://cushwake.cld.bz/2022-Global-Data-Center-Market-Comparison

[42] California ISO. 2016. Greenhouse Gas Emission Tracking Methodology. https://www.caiso.com/Documents/GreenhouseGasEmissionsTracking-Methodology.pdf

[43] California ISO. 2022. Managing Oversupply. https://www.caiso.com/informed/Pages/ManagingOversupply.aspx

[44] California ISO. 2022. Market Processes and Products. http://www.caiso.com/market/Pages/MarketProcesses.aspx

[45] Nicola Jones. 2018. How to Stop Data Centres from Gobbling up the World's Electricity. *Nature* (September 2018).

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

[47] Peter Judge. 2022. Dominion Energy admits it can't meet data center power demands in Virginia. https://www.datacenterdynamics.com/en/news/dominion-energy-admits-it-cant-meet-data-center-power-demands-in-virginia/

[48] Peter Judge. 2022. EirGrid pulls plug on 30 Irish data center projects. https://www.datacenterdynamics.com/en/news/eirgrid-pulls-plug-on-30-irish-data-center-projects/

[49] Peter Judge. 2022. Google and Microsoft join long term energy storage group. https://www.datacenterdynamics.com/en/news/google-and-microsoft-join-long-term-energy-storage-group/

[50] Kibaek Kim, Fan Yang, Victor Zavala, and Andrew A. Chien. 2016. Data Centers as Dispatchable Loads to Harness Stranded Power. *IEEE Transactions on Sustainable Energy* (2016). DOI 10.1109/TSTE.2016.2593607.

[51] Wook Hyun Kwon and Soo Hee Han. 2006. *Receding horizon control: model predictive control for state models.* Springer Science & Business Media.

[52] Kien Le, Ricardo Bianchini, Thu D. Nguyen, Ozlem Bilgir, and Margaret Martonosi. 2010. Capping the Brown Energy Consumption of Internet Services at Low Cost. In *Proceedings of the International Conference on Green Computing (GREENCOMP '10)*. IEEE Computer Society, USA, 3–14. https://doi.org/10.1109/GREENCOMP.2010.5598305

[53] Tan N Le, Zhenhua Liu, Yuan Chen, and Cullen Bash. 2016. Joint capacity planning and operational management for sustainable data centers and demand response. In *Proceedings of the Seventh International Conference on Future Energy Systems*. 1–12.

[54] Michelle Lewis. 2023. These big tech firms bought the most clean energy in 2022. Available from https://electrek.co/2023/02/09/big-tech-clean-energy-2022/.

[55] Liuzixuan Lin and Andrew A. Chien. 2020. *Characterizing Stranded Power in the ERCOT in Years 2012-2019: A Preliminary Report.* Technical Report TR-2020-06. University of Chicago.

[56] Liuzixuan Lin, Victor M Zavala, and Andrew A Chien. 2021. Evaluating Coupling Models for Cloud Datacenters and Power Grids. In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. 171–184.

[57] Minghong Lin, Adam Wierman, Lachlan LH Andrew, and Eno Thereska. 2012. Dynamic right-sizing for power-proportional data centers. *IEEE/ACM Transactions on Networking* 21, 5 (2012), 1378–1391.

[58] Julia Lindberg, Yasmine Abdennadher, Jiaqi Chen, Bernard C Lesieutre, and Line Roald. 2021. A Guide to Reducing Carbon Emissions through Data Center Geographical Load Shifting. In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. 430–436.

[59] Julia Lindberg, Line Roald, and Bernard Lesieutre. 2020. The Environmental Potential of Hyper-Scale Data Centers: Using Locational Marginal CO2 Emissions to Guide Geographical Load Shifting. In *Proceedings of the 54th Hawaii International Conference on System Sciences*. 3158.

[60] Zhenhua Liu, Yuan Chen, Cullen Bash, Adam Wierman, Daniel Gmach, Zhikui Wang, Manish Marwah, and Chris Hyser. 2012. Renewable and cooling aware workload management for sustainable data centers. In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*. 175–186.

[61] Z. Liu, M. Lin, A. Wierman, S. Low, and L. L. H. Andrew. 2015. Greening Geographical Load Balancing. *IEEE/ACM Transactions on Networking* 23, 2 (2015), 657–671.

[62] Zhenhua Liu, Adam Wierman, Yuan Chen, Benjamin Razon, and Niangjun Chen. 2013. Data center demand response: Avoiding the coincident peak via workload shifting and local generation. *Performance Evaluation* 70, 10 (2013), 770–791.

[63] Jianying Luo, Lei Rao, and Xue Liu. 2013. Temporal load balancing with service delay guarantees for data center energy cost optimization. *IEEE Transactions on Parallel and Distributed Systems* 25, 3 (2013), 775–784.

[64] Electricity Maps. 2022. Electricity Maps. https://app.electricitymaps.com/map

[65] Eric Masanet, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey. 2020. Recalibrating global data center energy-use estimates. *Science* 367, 6481 (2020), 984–986. https://doi.org/10.1126/science.aba3758 arXiv:https://science.sciencemag.org/content/367/6481/984.full.pdf

[66] Valérie Masson-Delmotte, Panmao Zhai, Hans-Otto Pörtner, Debra Roberts, Jim Skea, Priyadarshi R Shukla, Anna Pirani, W Moufouma-Okia, C Péan, R Pidcock, et al. 2018. Global warming of 1.5 C. *An IPCC Special Report on the impacts of global warming of* 1, 5 (2018).

[67] James McAnany. 2021. 2021 Demand Response Operations Markets Activity Report: January 2022. https://www.pjm.com/-/media/markets-ops/dsr/2021-demand-response-activity-report.ashx

[68] Ali Menati, Kiyeob Lee, and Le Xie. 2023. Modeling and analysis of utilizing cryptocurrency mining for demand flexibility in electric energy systems: A synthetic texas grid case study. *IEEE Transactions on Energy Markets, Policy and Regulation* (2023).

[69] Microsoft Azure. [n. d.]. Azure Global Datacenters. https://azure.microsoft.com/en-us/global-infrastructure/.

[70] Rich Miller. 2022. Dominion Energy Plans More Green Power for Virginia's Data Centers. https://datacenterfrontier.com/dominion-energy-plans-more-green-power-for-virginias-data-centers/

[71] Electric Reliability Council of Texas. [n. d.]. Large Flexible Load Task Force. https://www.ercot.com/committees/tac/lfltf.

[72] Anthony Papavasiliou and Shmuel S Oren. 2013. Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network. *Operations Research* 61, 3 (2013), 578–592.

[73] Brian Peccarelli. 2020. Three Ways COVID-19 is Accelerating Digital Transformation in Professional Services. (June 2020). https://bit.ly/34Qitb5, 37% growth.

[74] Sundar Pichai. 2019. Our biggest renewable energy purchase ever. https://blog.google/outreach-initiatives/sustainability/our-biggest-renewable-energy-purchase-ever/

[75] New York State Energy Planning Board. 2015. The Energy to Lead: 2015 New York State Energy Plan. http://energyplan.ny.gov/Plans/2015.aspx.

[76] SC Pryor, RJ Barthelmie, and TJ Shepherd. 2020. 20% of US electricity from wind will have limited impacts on system efficiency and regional climate. *Scientific reports* 10, 1 (2020), 1–14.

[77] Asfandyar Qureshi, Rick Weber, Hari Balakrishnan, John Guttag, and Bruce Maggs. 2009. Cutting the Electric Bill for Internet-Scale Systems. In *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication* (Barcelona, Spain) *(SIGCOMM '09)*. Association for Computing Machinery, New York, NY, USA, 123–134. https://doi.org/10.1145/1592568.1592584

[78] Ana Radovanovic, Ross Koningstein, Ian Schneider, Bokan Chen, Alexandre Duarte, Binz Roy, Diyue Xiao, Maya Haridasan, Patrick Hung, Nick Care, et al.

2021. Carbon-Aware Computing for Datacenters. *arXiv preprint arXiv:2106.11750* (2021).

[79] John Roach. 2021. Microsoft's virtual datacenter grounds 'the cloud' in reality. Microsoft to build 50 to 100 datacenters per year, https://news.microsoft.com/innovation-stories/microsofts-virtual-datacenter-grounds-the-cloud-in-reality/.

[80] Roy Schwartz, Jesse Dodge, Noah A. Smith, and Oren Etzioni. 2020. Green AI. *Commun. ACM* 63, 12 (nov 2020), 54–63. https://doi.org/10.1145/3381831

[81] Yuanyuan Shi, Bolun Xu, Baosen Zhang, and Di Wang. 2016. Leveraging energy storage to optimize data center electricity cost in emerging power markets. In *Proceedings of the Seventh International Conference on Future Energy Systems*. 1–13.

[82] Staff. 2020. COVID-19 Accelerates Cloud Adoption, Market to Reach $1 trillion, IDC. *Equipment FA News* (October 2020). https://www.equipmentfa.com/news/31459/covid-19-accelerates-cloud-adoption-market-to-reach-1t-idc.

[83] Grant L. Stewart, Gregory A. Koenig, Jingjing Liu, Anders Clausen, Sonja Klingert, and Natalie Bates. 2019. Grid Accommodation of Dynamic HPC Demand. In *Proceedings of the 48th International Conference on Parallel Processing: Workshops* (Kyoto, Japan) *(ICPP 2019)*. Association for Computing Machinery, New York, NY, USA, Article 9, 4 pages. https://doi.org/10.1145/3339186.3339214

[84] Qihang Sun, Shaolei Ren, Chuan Wu, and Zongpeng Li. 2016. An online incentive mechanism for emergency demand response in geo-distributed colocation data centers. In *Proceedings of the seventh international conference on future energy systems*. 1–13.

[85] Muhammad Tirmazi, Adam Barker, Nan Deng, Md E. Haque, Zhijing Gene Qin, Steven Hand, Mor Harchol-Balter, and John Wilkes. 2020. Borg: The next Generation. In *Proceedings of the Fifteenth European Conference on Computer Systems* (Heraklion, Greece) *(EuroSys '20)*. Association for Computing Machinery, New York, NY, USA, Article 30, 14 pages. https://doi.org/10.1145/3342195.3387517

[86] United Nations Framework Convention on Climate Change. 2015. Paris Climate Change Conference. http://unfccc.int/meetings/paris_nov_2015/meeting/8926.php.

[87] Rahul Urgaonkar, Bhuvan Urgaonkar, Michael J Neely, and Anand Sivasubramaniam. 2011. Optimal power cost management using stored energy in data centers. In *Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modeling of computer systems*. 221–232.

[88] Jonathan Vanian and Kif Leswing. 2023. ChatGPT and generative AI are booming, but the costs can be extraordinary. https://www.cnbc.com/2023/03/13/chatgpt-and-generative-ai-are-booming-but-at-a-very-expensive-price.html

[89] Markus Wacket. 2021. German parties agree on 2030 coal phase-out in coalition talks. https://www.reuters.com/business/cop/exclusive-germanys-government-in-waiting-agrees-phase-out-coal-by-2030-sources-2021-11-23/

[90] WattTime. 2022. Automated Emissions Reduction. https://www.watttime.org/solutions/automated-emissions-reduction-aer/

[91] Fan Yang and Andrew A Chien. 2016. ZCCloud: Exploring wasted green power for high-performance computing. In *2016 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*. IEEE, 1051–1060.

[92] Fan Yang and Andrew A. Chien. 2017. Large-scale and Extreme-Scale Computing with Stranded Green Power: Opportunities and Costs. *IEEE Transactions on Parallel and Distributed Systems* 29, 5 (December 2017).

[93] Lin Yang, Mohammad Hassan Hajiesmaili, Ramesh K. Sitaraman, Enrique Mallada, Wing Shing Wong, and Adam Wierman. 2019. Online Inventory Management with Application to Energy Procurement in Data Centers. *CoRR* abs/1901.04372 (2019). arXiv:1901.04372 http://arxiv.org/abs/1901.04372

[94] Chaojie Zhang. 2022. *Eliminating the Capacity Variation Penalty for Cloud Resource Management*. Ph. D. Dissertation. The University of Chicago, Chicago, IL, USA. Advisor(s) Chien, Andrew A.

[95] Chaojie Zhang and Andrew A Chien. 2021. Scheduling Challenges for Variable Capacity Resources. In *Workshop on Job Scheduling for Parallel Processing (JSSPP)*.

[96] Linquan Zhang, Shaolei Ren, Chuan Wu, and Zongpeng Li. 2015. A truthful incentive mechanism for emergency demand response in colocation data centers. In *2015 IEEE Conference on Computer Communications (INFOCOM)*. IEEE, 2632–2640.

[97] Weiqi Zhang, Line A Roald, Andrew A Chien, John R Birge, and Victor M Zavala. 2020. Flexibility from networks of data centers: A market clearing formulation with virtual links. *Electric Power Systems Research* 189 (2020), 106723.

[98] Yanwei Zhang, Yefu Wang, and Xiaorui Wang. 2012. Electricity bill capping for cloud-scale data centers that impact the power markets. In *2012 41st International Conference on Parallel Processing*. IEEE, 440–449.

[99] Jiajia Zheng, Andrew A. Chien, and Sangwon Suh. 2020. Mitigating Curtailment and Carbon Emissions through Load Migration between Data Centers. *Joule* (October 2020). https://doi.org/10.1016/j.joule.2020.08.001

[100] Zhi Zhou, Fangming Liu, Shutong Chen, and Zongpeng Li. 2018. A truthful and efficient incentive mechanism for demand response in green datacenters. *IEEE Transactions on Parallel and Distributed Systems* 31, 1 (2018), 1–15.

[101] Zhi Zhou, Fangming Liu, Bo Li, Baochun Li, Hai Jin, Ruolan Zou, and Zhiyong Liu. 2014. Fuel cell generation in geo-distributed cloud services: A quantitative study. In *2014 IEEE 34th International Conference on Distributed Computing Systems*. IEEE, 52–61.

[102] Zhi Zhou, Fangming Liu, Ruolan Zou, Jiangchuan Liu, Hong Xu, and Hai Jin. 2016. Carbon-Aware Online Control of Geo-Distributed Cloud Services. *IEEE Transactions on Parallel and Distributed Systems* 27, 9 (2016), 2506–2519.

A DIRECT-CURRENT OPTIMAL POWER FLOW FORMULATION

We model the grid operation using the direct-current optimal power flow (DC-OPF) model. Notations in this model are listed below:

Table 3: DC-OPF Notation: Sets

Notation	Description	Notation	Description
$\mathcal{D}$ ($\mathcal{D}_n$)	Demand loads (at bus n)	$\mathcal{G}$ ($\mathcal{G}_n$)	Generators (at bus n)
$\mathcal{I}$ ($\mathcal{I}_n$)	Import points (at bus n)	$\mathcal{L}$	Transmission lines
$\mathcal{L}_n^+/\mathcal{L}_n^-$	Transmission lines to/from bus n	$\mathcal{N}$	Buses
$\mathcal{R}$ ($\mathcal{R}_n$)	Renewable generators (at bus n)	$\mathcal{T}$	Time periods
$\mathcal{W}$ ($\mathcal{W}_n$)	Wind farms (at bus n)	$DC(DC_n)$	Datacenters (at bus n)

Table 4: DC-OPF Notation: Parameters

Notation	Description	Notation	Description
B_l	Susceptance of transmission line l	C_i	Generation cost of generator i
C_j^d	Load-shedding penalty at load j	C_i^w	Curtailment penalty at wind farm i
C_i^m	Curtailment penalty at import point i	C_i^r	Curtailment penalty at renewable i
$D_{j,t}$	Demand load of consumer j at time t	F_l^{max}	Maximum power flow of transmission line l
$M_{i,t}$	Power production of import i at time t	P_i^{max}	Maximum power output of generator i
$R_{i,t}$	Power production of renewable i at time t	RU_i	Ramp-up limit of generator i
RD_i	Ramp-down limit of generator i	$W_{w,t}$	Power from wind farm w at time t
$\Theta_{n,t}^{min}$	Minimum phase angle at bus n at time t	$\Theta_{n,t}^{max}$	Maximum phase angle at bus n at time t

Table 5: DC-OPF Notation: Decision Variables

Notation	Description	Notation	Description
$d_{j,t}$	Load shedding at load j at time t	$f_{l,t}$	Power flow of line l at time t
$m_{i,t}$	Curtailment at import i at time t	$p_{i,t}$	Power from generator i at time t
$r_{i,t}$	Curtailment at renewable i at time t	$w_{i,t}$	Curtailment at wind farm i at time t
$\theta_{n,t}$	Phase angle at bus n at time t		

Datacenter capacity levels $cap_{i,t}$ are external decisions by the datacenters or coordinator(s). The power grid solves the DC-OPF model (one-day time horizon with hourly intervals in our studies), minimizing the dispatch cost (6a) that consists of generation cost, load shedding penalty, and curtailment penalties:

$$\min \quad \sum_{t\in\mathcal{T}}\left(\sum_{i\in\mathcal{G}} C_i p_{i,t} + \sum_{j\in\mathcal{D}} C_j^d d_{j,t} + \sum_{i\in\mathcal{I}} C_i^m m_{i,t} + \sum_{i\in\mathcal{W}} C_i^w w_{i,t} + \sum_{i\in\mathcal{R}} C_i^r r_{i,t}\right) \tag{6a}$$

$$\text{s.t.} \quad \sum_{l\in\mathcal{L}_n^+} f_{l,t} - \sum_{l\in\mathcal{L}_n^-} f_{l,t} + \sum_{i\in\mathcal{G}_n} p_{i,t} + \sum_{i\in\mathcal{I}_n}(M_{i,t}-m_{i,t}) + \sum_{i\in\mathcal{W}_n}(W_{i,t}-w_{i,t}) + \sum_{i\in\mathcal{R}_n}(R_{i,t}-r_{i,t}) = \sum_{j\in\mathcal{D}_n}(D_{j,t}-d_{j,t}) + \sum_{i\in DC_n} cap_{i,t}, \quad \forall n\in\mathcal{N}, t\in\mathcal{T}, \tag{6b}$$

$$f_{l,t} = B_l(\theta_{n,t}-\theta_{m,t}), \quad \forall l=(m,n)\in\mathcal{L}, t\in\mathcal{T}, \tag{6c}$$

$$-F_l^{max} \le f_{l,t} \le F_l^{max}, \quad \forall l\in\mathcal{L}, t\in\mathcal{T}, \tag{6d}$$

$$\Theta_n^{min} \le \theta_{n,t} \le \Theta_n^{max} \quad \forall n\in\mathcal{N}, t\in\mathcal{T}, \tag{6e}$$

$$-RD_i \le p_{i,t} - p_{i,t-1} \le RU_i, \quad \forall i\in\mathcal{G}, t\in\mathcal{T}, \tag{6f}$$

$$0 \le p_{i,t} \le P_i^{max}, \quad \forall i\in\mathcal{G}, t\in\mathcal{T}, \tag{6g}$$

$$0 \le d_{j,t} \le D_{j,t}, \quad \forall j\in\mathcal{D}, t\in\mathcal{T}, \tag{6h}$$

$$0 \le m_{i,t} \le M_{j,t}, \quad \forall i\in\mathcal{I}, t\in\mathcal{T}, \tag{6i}$$

$$0 \le w_{i,t} \le W_{j,t}, \quad \forall i\in\mathcal{W}, t\in\mathcal{T}, \tag{6j}$$

$$0 \le r_{i,t} \le R_{j,t}, \quad \forall i\in\mathcal{R}, t\in\mathcal{T}. \tag{6k}$$

The generation sources include conventional thermal power plants (e.g. gas, nuclear, coal), non-wind renewables (e.g. hydro), imports, and wind power plants. Due to long-term commitments for imports or goal of reducing carbon emissions with renewables, the imports and renewables are non-dispatchable as in [50] but can be curtailed at the cost of C_i^m and C_i^r \$/MWh (C_i^w \$/MWh for wind) respectively. Sometimes the power supply may not meet the demand, and each unit of load shedding (not served load) is at the cost of value of lost load (VOLL) C_j^d. In this paper, the unit generation cost is 1/2/4 \$/MWh for nuclear/coal/gas. The penalties are 500/100/1,000 \$/MWh for import/wind/non-wind renewables curtailment, 1,000 \$/MWh for load shedding.

The constraints are typical for DC-OPF. Constraint 6b represents the balance constraint at each bus, whose associated dual value is the locational marginal price (LMPrice) at that bus, indicating marginal cost of adding 1 MW load at a specific location, so the price can go negative/high when curtailment/load shedding happens. Constraints 6c–6e represent how the power flow (6c) is determined given the line capacity (6d) and phase angle (6e) limits. Constraint 6f limits conventional power plants' rate of ramping up/down generation. Constraints 6g–6k bound the conventional generation, load shedding, and curtailments (6i–6k) respectively.

B FUEL CARBON EMISSION RATES

Below are the fuel carbon emission rates (carbon emissions per MWh energy generated from that fuel) we use to calculate carbon emissions:

Table 6: Carbon Emission Rates of Different Fuels [6, 42]

Generation Type	Carbon Emission Rate (kg CO_2/MWh)
Coal	895.2
Natural Gas	388.9
Oil	877.6
Dual-fuel	633.3
Nuclear	0
Geothermal	107.6
Biomass	0
Hydro	0
Wind	0
Import	428