

Distribution and Management of Datacenter Load Decoupling

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

The exploding power consumption of AI and cloud datacenters (DCs) intensifies the long-standing concerns about their carbon footprint, especially because DCs' need for constant power clashes with volatile renewable generation needed for grid decarbonization. DC flexibility (a.k.a. load adaptation) is a key to reducing DC carbon emissions by improving grid renewable absorption.

DC flexibility can be created, without disturbing datacenter capacity by *decoupling* a datacenter's power capacity and grid load with a collection of energy resources. Because decoupling can be costly, we study how to best distribute and manage decoupling to maximize benefits for all. Key considerations include site variation and datacenter-grid cooperation.

We first define and compute the power and energy needs of datacenter load decoupling, and then we evaluate designed distribution and management approaches. Evaluation shows that optimized distribution can deliver >98% of the potential grid carbon reduction with 70% of the total decoupling need. For management, DC-grid cooperation (2-way sharing and control vs. 1-way info sharing) enables 1.4x grid carbon reduction. Finally, we show that decoupling may be economically viable, as on average datacenters can get power cost and carbon emissions benefits greater than their local costs of decoupling. However, skew across sites suggests grid intervention may be required.

CCS Concepts

• **Applied computing** → **Data centers**; • **Hardware** → **Power and energy**; • **Social and professional topics** → **Sustainability**.

Keywords

Load adaptation, Datacenter energy management, Sustainable computing

1 Introduction

Datacenters are one of the fastest-growing sectors of power consumption in many regions around the world [38]. In the U.S., from 2018 to 2022, datacenters' power consumption grew 17% annually, driven by rapid cloud growth. However, the annual growth has doubled since 2022, with projected growth of 27%/year through 2028 and beyond [77]. Similar growth is projected worldwide [16]. This growth is fueled by expanding AI applications and cloud business, reflected in hyperscalers' exploding capital expenditures (CapEx). The total CapEx of the 11 largest hyperscalers in the world has grown to \$392B in 2025. This number is a record high that exceeds the annual investments in 2022 and 2023 combined [27].

This accelerated growth intensifies the long-standing concerns about datacenters' growing power consumption and associated carbon footprint [1, 60]. In addition to the growth, a root cause of

the concerns is the conflict between datacenters' need for constant power and the variation inherent in volatile renewable generation needed for grid decarbonization. To avoid decarbonization slow-down and to protect grid reliability, some grid operators delay approval or even suspend new datacenter projects [47, 48]. In some other cases, the grid delay the decommission of coal power plants or propose building new gas power plants to support datacenter growth [59, 75, 80], which provide reliable power supply at low cost but clearly destructive to grid decarbonization. Overall, this conflict put either datacenter growth or grid decarbonization (and greening compute along with it) at risk.

Need for Datacenter Flexibility. Hyperscalers have sustainability commitments such as "net-zero by 2030", which require a significant decrease in datacenter carbon emissions. Although they are purchasing renewable power through power purchase agreement (PPA) to "offset" their power consumption, the increased need of new gas generators in the grids and the 15–20% annual growth of Scope 2 carbon emissions reported [18, 31] reflect datacenters' heavy reliance on fossil-fuel generation. Datacenter flexibility to match load to renewable generation is the key to achieving DC decarbonization goals. In power grids today where many include 10–40% renewables with growth to 40–90% by 2030 [39], significant renewable generation is wasted ("curtailed") or sold at zero or negative prices [9, 17, 51, 97]. For example, curtailment in CAISO (California) was 3.4 TWh in 2024, having grown 40%/year in 2015–2023 [9]. Further, gas generation powers datacenters at night, or when wind generation is insufficient. Datacenter flexibility can reduce both the curtailment and the use of fossil-fuel generation. Finally, the promise of datacenter flexibility to reduce DC carbon emissions has been shown by extensive research [19, 21, 30, 36, 52, 57, 58, 82, 84]. However, because computing workloads traditionally require constant power, the adoption of datacenter flexibility in industrial practice is rare or on a limited scale.

Need for Decoupling. Cloud datacenters' need for constant power is inherent in quality of service (QoS) commitments (e.g. 99.999% availability) and business pressure for maximum utilization. Varying datacenter power capacity to match volatile renewable generation harms datacenter productivity [95, 100], which perhaps explains the limited flexibility in Google's Carbon-aware Computing [72]. Furthermore, traditional load flexibility incentives including time-varying power rate and incentive payments are not attractive to cloud datacenters as power cost is only a small fraction relative to the revenue of cloud services [4, 5].

What if we could achieve both stable datacenter capacity and high renewable power absorption? That would maintain high IT efficiency while enabling the power grid to decarbonize. For backup, many datacenters deploy energy resources such as generators and batteries [49, 81, 85]. Such energy resources can

be used not just in emergencies, but more generally to *decouple* datacenter’s power capacity from its grid load. This enables IT operation to be unaffected while the grid load is flexed to avoid carbon emissions. **But how much energy resource should be deployed and where?** With complex grid structure and dynamics, deciding whether additional energy resources are beneficial locally is difficult. It’s even more challenging to understand if they would be more valuable at another datacenter site. **Beyond distribution, intelligent management is another challenge.** As large loads up to gigawatts [62] and collectively accounting for 20–30% of grid load [24, 33], datacenters that flex grid load to match renewable generation in the grid can overshift their load, disrupting grid dispatch [52]. Such management would require intelligent coordination between DCs and the grid to achieve maximum benefits.

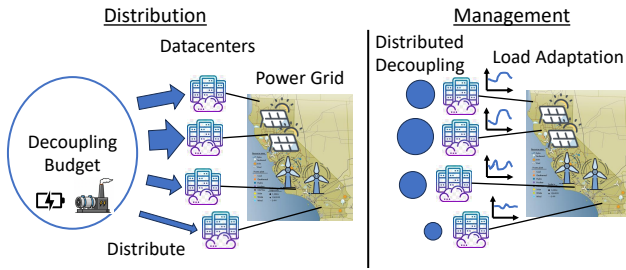


Figure 1: Two Phases of Datacenter Load Decoupling: (1) provision energy resources for decoupling at various datacenter sites; (2) manage decoupling to adapt datacenter grid load to grid dynamics.

With the goal of maximizing DC and grid carbon reduction at low cost, we study the distribution and management of decoupling energy resources. We first analyze grid dynamics and ideal load profiles to understand decoupling requirements. Then, as shown in Figure 1, we distribute a total decoupling budget across datacenters for maximum potential benefits. Finally, we consider coordination and management approaches that exploit distributed decoupling efficiently for carbon emissions reduction. Specific contributions include:

- Definition and computation of datacenter decoupling power and energy needs.
- Given a global decoupling budget, define a method that distributes it for high effectiveness. Specifically, our stochastic optimization approach delivers 98–100% of potential grid carbon reduction with only 70% of the total decoupling need.
- Design and evaluation of decoupling management approaches. We propose PS-GridScale, a new 2-way sharing and control approach, comparing it with 1-way information sharing (PlanShare) and grid-optimal control approaches. PS-GridScale achieves up to 1.4x further grid carbon reduction vs. PlanShare and captures 84–90% of the potential benefits.
- Financial analysis that shows decoupling may be economically viable. On average, DCs gain power cost and carbon emissions benefits greater than their local costs of decoupling. However, the skew across sites suggests some grid

intervention may be required for implementation. In addition, we show power grids may gain more decarbonization from adding decoupling than building additional renewable generation.

The remainder of the paper is organized as follows. Section 2 introduces the background of need for datacenter flexibility and decoupling. Section 3 motivates the study of distribution and management of decoupling and summarizes our approach. And then we elaborate the design of distribution and management approaches in Section 4, followed by corresponding evaluation in Section 5. Finally, we discuss the related work and conclude.

2 Background

2.1 Power Grid Decarbonization

In response to global warming, power grid decarbonization, which replaces fossil-fuel generation with carbon-free energy, is taking place all over the world [6, 14, 20, 25, 26]. For example, aiming at 80% renewables in electricity by 2030, Germany reached a record share of 62.7% in 2024 [8]. China, the world’s largest electricity producer, built 357 GW of solar and wind generation capacity in 2024 [67]. After seeing 100% load supported by renewables for the first time in 2022, California now regularly sees this phenomenon with annual average renewable fraction of 39% [45]. Globally, IEA projects that renewable capacity is expected to increase over 5,520 GW during 2024–2030, 2.6 times more than the deployment of the previous six years (2017–2023) [39].

The growth of renewable generation capacity is dominated by wind and solar that depend on natural dynamics of wind and sun. As the fraction of renewable generation increases, the challenge of balancing demand and generation increases: when renewable generation is low, power outage can happen; when renewable generation is high, it may be wasted (so-called “curtailment”) [9, 10, 51, 97] or produce negative prices in the wholesale power market [17, 83]. Supply-side solutions to balancing include energy storage [44], generation capacity overprovisioning [34, 71], and peaking power plants. However, relying solely on these solutions is costly because of the long-tail statistics of wind and solar. Demand-side solutions that adapt load to grid dynamics can significantly reduce infrastructure cost [28].

2.2 Datacenters as Growing Fixed Loads

Growth. Datacenters are growing in both size and share of grid load. Campuses of hundreds of megawatts are common today, and giants of multiple gigawatts are on the way [62]. Growth as a fraction of grid load is projected to continue at current hotspots, including Northern Virginia (25% in 2023, 49% in 2034 projected) and Ireland (24% in 2024, 30% in 2032 projected). High datacenter penetration is spreading. For example, Texas and Arizona expect 16% and 17% grid load from datacenters in 2030 respectively [40, 66]. With this scale and fraction, it’s important for the grids to consider datacenters when making policies and planning resources.

Fixed Grid Load. Datacenters are expensive capital assets for which hyperscalers seek to maximize utilization [87]. The corresponding job scheduling practice (e.g. oversubscription [73]) and

service models (e.g. constant-rate VMs) don't tolerate power variation, as it can cause job termination or increased service latency [95, 100]. And despite a decade of effort on energy-proportionality, inflexible idle server power still corresponds to 60% of energy consumption in commercial clouds [76]. Together, these result in cloud datacenters' highly fixed and inflexible grid load.

2.3 Realizing Datacenter Flexibility

Datacenter grid load flexibility is the key to reconciling grid decarbonization and datacenter growth as reflected in a major EPRI DCFlex initiative involving all of the leading hyperscalers and grid companies in the U.S. [41]. Generally, grid load flexibility can be realized by power use behavior changes and behind-the-meter energy storage [28], which correspond to IT operation and decoupling for datacenters. Many academic researchers have explored flexing datacenter grid load through dynamic resource scaling (e.g. toggle servers into or out of power saving mode [55], software-based autoscaling [36]), and batch workload scheduling [84, 93]. However, it's hard to adopt these approaches in commercial cloud datacenters due to cloud providers' concerns about service impact and associated revenue loss. Another approach is flexing of datacenter capabilities including cooling and DVFS [11, 29]. One promising approach is load decoupling with energy resources as it avoids impacts on IT operation. Because of the limitations of UPS energy storage (limited capacity [91]) and diesel backup generators (carbon-heavy generation) currently deployed in datacenters, load decoupling for carbon reduction would incur additional cost. However, hyperscalers' recent actions of adding generator or energy storage to datacenters in response to requirements of grid operators [49, 81, 89] suggest it is a practical solution.

3 Problem and Approach

Problem. In power grids with significant volatile renewable generation, flexing datacenters' grid load to match renewable generation can reduce grid carbon emissions, dispatch cost, etc. This can mitigate the grid challenges that DC growth creates. Figure 2 illustrates a cost-benefit space. At the upper left are fixed load DCs (current practice). At the lower right are flexible load DCs, controlled by grid dispatch to maximize grid welfare. Modeling results show that flexibility in DCs can reduce grid carbon emissions by 10% in this wind-dominated grid. Achieving this benefit requires significant flexibility in DC load. One measure of daily flexibility, the total maximum daily energy deficit (defined in Section 4.1), is 86,400 MWh (21% of daily power consumption) to achieve it.

Achieving grid load variation simultaneously with stable DC capacity, e.g. decoupling, is difficult and can be expensive. Decoupling of DC grid load and power capacity can be achieved with a collection of co-located energy resources (e.g. energy storage, generators). If DC capacity variation is allowed, it can be harmful to DC productivity [100], which incurs high direct and indirect costs [42]. Therefore, We focus on the decoupling approach. Decoupling allows compute performance unaffected while the grid load is flexed, but it also incurs costs (e.g. equipment, operation) additional to current datacenter infrastructure. Naturally, a key question for load decoupling is: **are the benefits worth the costs? and does**

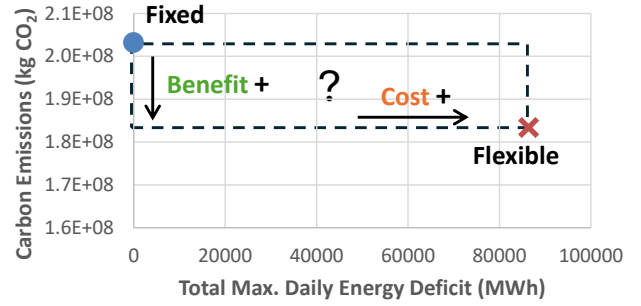


Figure 2: Datacenter flexibility reduces grid carbon emissions but realizing has a cost, forming a trade-off space to explore.

the answer differ from DC and grid perspectives? Subsidiary research questions include:

- What is the cost-benefit trade-off?
- How to distribute a total decoupling budget across datacenters for maximum potential benefits?
- How to efficiently manage the distributed decoupling capability?
- How to attribute the benefits to incentivize support for load adaptation?
- Does the grid generation composition affect these answers?

Approach. To understand the need for decoupling for load flexibility, we first define the key dimensions of load profiles that capture the power and energy capacity required. These dimensions characterize the decoupling needs to maximize grid benefits, and we use them as starting points in the trade-off space (e.g. the cross in Figure 2).

With the formalized needs, designing DC load decoupling has two phases: distribution and management. First, distribution decides the static decoupling capacity at each datacenter given a total budget. Second, management controls the distributed capacity to flex datacenter grid load. These two phases together produce the costs (capital, operational) and benefits (operational) of load decoupling. Varying the total decoupling budget frames the trade-off.

We explore the design space for distribution and management of datacenter load decoupling with the goal of achieving high datacenter and grid benefits at low cost. In the following sections, we proceed as follows:

- (1) Analyze the power and energy needs of decoupling given targeted datacenter grid load and capacity profiles.
- (2) Propose several distribution and management approaches for decoupling with desirable properties (e.g. effectiveness, datacenter autonomy).
- (3) Evaluate their costs and benefits from the perspectives of both datacenters and the grid.

4 Designing Datacenter Load Decoupling

4.1 What is the Local Decoupling Need?

Datacenter operators typically aim to achieve constant, high resource capacity (machines and power) utilization. This maximizes

capital efficiency, and productive output for very expensive datacenters [27, 79]. The corresponding workload management, along with hardware constraints such as poor server energy proportionality, produces stable (nearly constant) datacenter grid loads. If to help the grid, datacenter power capacity were varied with grid dynamics, that would in many cases harm compute efficiency and associated service revenue [95, 100]. We consider an alternative, decoupling the datacenter power capacity from its grid load with energy resources (e.g. generators and energy storage). This approach enables the best of both worlds, flexible grid load and constant datacenter power capacity, but at a cost.

We characterize local decoupling need for a datacenter by studying the power (instantaneous) differences between targeted grid load ($gridLoad_t$) and datacenter power capacity ($DCPower_t$). We also consider the energy (accumulated power) over time. Formally, for any time point t in interval $[t_1, t_2]$ ¹:

Power surplus: $decPow^+(t) = (gridLoad_t - DCPower_t)^+$

Power deficit: $decPow^-(t) = (DCPower_t - gridLoad_t)^+$

As the power difference accumulates, we have:

Energy surplus: $decEn^+(t_1, t_2) = \sum_{t=t_1}^{t_2} (gridLoad_t - DCPower_t)^+$

Energy deficit: $decEn^-(t_1, t_2) = \sum_{t=t_1}^{t_2} (DCPower_t - gridLoad_t)^+$

Net energy: $decEn_{net}(t_1, t_2) = decEn^+(t_1, t_2) - decEn^-(t_1, t_2)$

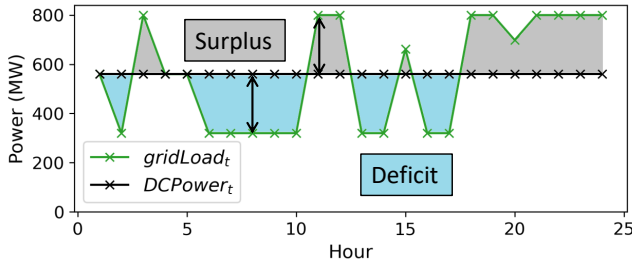


Figure 3: Flexed datacenter grid load $gridLoad_t$ produces surplus/deficit in power/energy relative to current datacenter operation practice represented by $DCPower_t$.

Aligning these terms with the power grid's daily dispatch and the targeted constant $DCPower$ profile ($DCPower_t = DCPower_{avg}$, $\forall t$), over 24 hours on day d (Figure 3):

$$decEn_d^+ = \sum_{t=1}^{24} (gridLoad_t - DCPower_{avg})^+$$

$$decEn_d^- = \sum_{t=1}^{24} (DCPower_{avg} - gridLoad_t)^+$$

¹ $(x)^+ = x$ if $x \geq 0$, else $(x)^+ = 0$.

Specifically, we consider flexible $gridLoad_t$ adapted to grid dynamics with $DCPower_{avg}$ as average load throughout a day, which benefits grid reliability and decarbonization [28]. Formally,

$$util_{min} \leq gridLoad_t / DCPower_{max} \leq util_{max} \quad (1a)$$

$$decPEn_{net}(0, t) \leq 0, \quad decPEn_{net}(0, 24) = 0 \quad (1b)$$

Considering the energy resources that decouples constant $DCPower$ from this flexible $gridLoad$, Eq. 1b enforces a circular boundary constraint across days. In the long term, the local decoupling need is defined in terms of maximum daily power deficit (aggregated power) and energy deficit (energy capacity) respectively:

$$decPow_{max}^- = DCPower_{max} \cdot (util_{avg} - util_{min})$$

$$decPEn_{max}^- = \max(\{decPEn_d^-\})$$

The decoupling need is a key determinant of the capital costs of energy resources, such as the fuel storage for a generator or the energy capacity of a battery. In the following sections, we explore how to distribute a total budget that constrains $\sum_{i \in DC} decPEn_{i,max}^-$ and manage local decoupling capacity ($decPEn_{i,max}^-$).

4.2 How to Distribute Decoupling?

We consider a global “total decoupling” budget, and ask how to distribute it most effectively? We assume that the decoupling, tied to physical resources, must be done statically, and consider two distribution methods, prioritizing fairness or efficiency.

Even distribution (EvenDist) evenly distributes the total budget to datacenters (or proportional to their power load). This simple method treats different datacenters equally in allocating decoupling. This is similar to EirGrid's requirement that new datacenters need to bring generation or storage resources equivalent to their load for grid connection [33].

Grid-optimized distribution (OptDist) distributes decoupling in a fashion that optimizes grid metrics such as dispatch cost. This approach takes into account grid structure (e.g. transmission, generation location) as well as datacenter location. For example, distributing more decoupling to datacenters close to renewable generators or experiencing severe transmission congestion can be more beneficial to the grid. To find OptDist distributions, we formulate a stochastic optimization problem. Formally,

$$\min \sum_{d \in S} w_d \cdot dispatchCost_s \quad (2a)$$

$$\text{s.t.} \quad \sum_{i \in DC} decPEn_{i,max}^- \leq totalDecPEn \quad (2b)$$

$$decPEn_{i,d}^- \leq decPEn_{i,max}^-, \quad \forall d \quad (2c)$$

$$gridLoad \text{ flexibility constraints (1a), (1b), } \forall d$$

$$\text{grid DC-OPF constraints, } \forall d$$

The objective is to minimize the weighted average grid dispatch cost, a metric that reflects overall grid welfare (defined in Section 5.1), over a set of scenarios S covering season and weekday/weekend variation. Constraint 2b distributes the total budget $totalDecPEn$ to different datacenters as a static property across the scenarios. For each day $d \in S$, datacenter load adaptation produces energy deficit not more than the distributed limit (2c). This formulation is a special case of the optimal grid resource planning, where the resources are only added to datacenter sites to directly mitigate the

negative impacts of datacenter growth. More general placement arrangements are possible [46, 96].

By varying the total budget $totalDecpEn$, we can explore the trade-off between decoupling capacity and benefits.

Note that another obvious way to distribute decoupling capacity would be in proportion to maximum local need. We revisit this question in the discussion of Figure 5.

4.3 How to Manage Distributed Decoupling?

If decoupling is owned or paid for by datacenters, then the operators may prefer to preserve some management autonomy. This is different from GridCtrl scenario considered where the power grid controls the decoupling resources. We frame two additional management approaches that preserve autonomy and also GridCtrl. Where datacenters have management autonomy, cooperation with the grid is also considered as it's important for effective grid carbon reduction [52].

DC-adapt with 1-way Info Sharing. This approach extends classical local, selfish load adaptation, sharing information about future datacenter load with the grid. This approach preserves the autonomy of datacenters in $gridLoad_t$ while enabling grid dispatch to exploit the load information for optimization. There are a diverse set of grid metrics (e.g. power price, carbon intensity) [56] which datacenters can adapt grid load to based on the objective (e.g. minimizing power cost, carbon emissions), availability, and effectiveness. Specifically, we evaluate **PlanShare** [52] that makes a 24-hour load plan adapted to locational marginal price (LMP)²—a widely usable metric correlated with local carbon intensity—and shares that load plan with the grid day-ahead. Formally, for a datacenter's operation in a day, $\{gridLoad_t\}$ is decided by:

$$\min \sum_{t=1}^{24} LMP_t \cdot gridLoad_t \quad (3a)$$

$$\text{s.t. } |gridLoad_t - gridLoad_{t-1}| \leq stepSize, 2 \leq t \leq 24 \quad (3b)$$

$$decpEn_d^- \leq decpEn_{max}^- \quad (3c)$$

Constraints (1a), (1b)

Datacenters optimize their own power cost, and their load plan sharing enables grid-wide benefits. Constraint 3b limits large load changes that could harm the grid, where $stepSize$ is tuned based on grid configurations. We add constraint 3c to ensure the load shape meets the decoupling energy limit.

2-way Info Sharing and Control. Load variation in grid load profiles fully determined by datacenters is subject to local view and can cause grid problems even with PlanShare. For example, LMP only indicates the type of generator that will be dispatched next if load increases but doesn't indicate how much load increase it can support, which is a common issue with marginal metrics. Synchronized behaviors of increasing load due to correlated LMPs across datacenter sites can oversubscribe the renewable curtailment. In contrast, the global view and control of grid dispatch enable it to avoid such problems. We design a new coordination approach—**P(lan)S(hare)-GridScale**—that enables the grid to modify the load

plans proposed by datacenters. First, datacenters make 24-hour load plans as PlanShare. And then the grid imposes scale factors $\{\alpha_{i,t}\}$ on the load deviation from constant load at different datacenters to further reduce grid dispatch cost. Formally, let $gridLoad'_{i,t}$ denote datacenter i 's proposed grid load at time t , $gridLoad_{i,t}$ denote the load finalized by the grid. Decision variables $\{\alpha_{i,t}\}$ and the following constraints ($t = 1, \dots, 24, \forall i \in DC$) are added to the grid dispatch:

$$gridLoad_{i,t} = DCPower_{i,avg} + \alpha_{i,t}(gridLoad'_{i,t} - DCPower_{i,avg}) \quad (4a)$$

$$0 \leq \alpha_{i,t} \leq 1 \quad (4b)$$

PS-GridScale combines datacenter autonomy and global control of the grid. A property of PS-GridScale favorable to datacenters is that the magnitude of decoupling power after grid modification doesn't exceed the initial proposal.

Grid-controlled Adaptation. Datacenter load levels are decided by the grid dispatch, respecting the flexibility constraints (e.g. dynamic range of $gridLoad_t$). This is an idealized approach, not practiced in today's power grids.³ In this paper, we focus on a variant (**GridCtrl**) targeting the mismatch between datacenters' fixed grid load and variable renewables, which regularly adjusts DCs' grid load to maximize grid benefits:

$$\min \text{dispatchCost} \quad (5a)$$

$$\text{s.t. } decpEn_d^- \leq decpEn_{i,max}^- \quad (5b)$$

Constraints (1a), (1b)

grid DC-OPF constraints

GridCtrl decides datacenters' $\{gridLoad_{i,t}\}$ as additional decision variables in the DC-OPF problem (described in Section 5.1). This approach provides reference of maximum grid benefits but might incur high costs on the datacenter side. Similar pilot projects such as ERCOT's "controllable load resources" program [65] are ongoing in power grids to facilitate grid connection of large loads.

5 Evaluation

5.1 Experimental Setup

5.1.1 Grid Simulation. Our grid simulation is based on a reduced California power grid (CAISO) model from [69] with 225 buses (more details in Appendix A). We update the generation cost of conventional generators with the latest EIA annual data [3], and adding wind generation (scale existing sites) and solar generation (add new sites [43, 98]) as appropriate to create the grid types discussed below.

Because renewable penetration⁴ has already exceeded 60% in Germany, Peru, and Canada (and many others), and is the goal of many grids in the 2030s [22, 33, 39], we assume that level of renewables and vary mix. We study three grid types: (1) 60% wind, 0% solar (**wind-dominated**); (2) 30% wind, 30% solar (**balanced**); (3) 0% wind, 60% solar (**solar-dominated**). Examples of these include EirGrid in Ireland is wind-dominated, ERCOT in Texas is projected to be balanced, and CAISO in California is solar-dominated.

²Other grid metrics such as locational marginal carbon intensity are also usable as information availability evolves.

³The only time this happens to a limited degree is during grid emergencies to avoid grid collapse [15].

⁴Renewable penetration is defined as the ratio of renewable generation to grid demand.

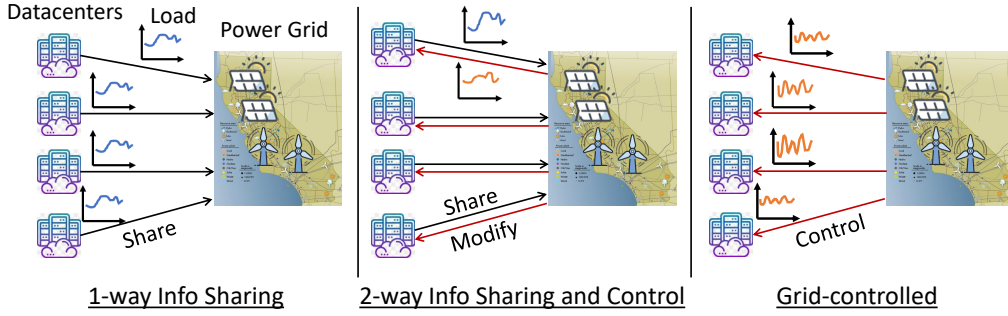


Figure 4: Decoupling Management Approaches varying in Datacenter Autonomy. PlanShare, PS-GridScale, and GridCtrl belong to the three categories respectively.

The grid model is associated with 8 profiles of base load, import, and renewable generation (wind and solar excluded) corresponding to weekday and weekend day of four seasons. For each season, we use 25 wind and solar generation profiles to model the day-to-day generation variation.

30 datacenters, each with $DCPower_{max} = 800$ MW and $avgUtil = 70\%$ are added at 30 random buses in the grid. The aggregated 16.8 GW average load corresponds to 38% of grid load, matching real-world grid projections for 2030–2035 [24, 33, 86].

With the attributes or profiles of grid entities as input, the direct-current optimal power flow (DC-OPF) problem [37] is solved by the grid for generator dispatch in a day. DC-OPF minimizes the dispatch cost (generation cost + curtailment/load shedding penalties), subject to balancing, transmission, and generator constraints. We attach the full formulation in Appendix A for reviewer convenience.

Table 1: Summary of Power Grid and Datacenter Setup

Power Grid (based on CAISO topology)	
Base Load	27, 289 MW on average
(Wind%, Solar%)	(60, 0), (30, 30), (0, 60)
Datacenters (30 sites)	
$DCPower_{max}$	800 MW
$DCPower_{avg}$	560 MW ($util_{avg} = 70\%$)
$[util_{min}, util_{max}]$	[40%, 100%]

5.1.2 Decoupling Settings. The dynamic range for $gridLoad_t$ is [40%, 100%], enabling large potential benefits. This range defines a maximum power deficit/surplus of 240 MW. For total decoupling budget, $\sum_{i \in DC} decpEn_{i,max}^-$, we consider a range from 0 to 100% of the maximum need, exploring the full range.

We construct a dataset of 200 days (8 day types * 25 wind/solar scenarios), splitting it into 4:1 for creating the optimized distribution (OptDist) and evaluating management respectively.

5.1.3 Implementation. The distribution and management of load decoupling and the grid DC-OPF are implemented as linear programs (LPs) with Julia/JuMP [23]. We solve the LPs with Gurobi solver [35]. The largest LP, stochastic optimization for OptDist takes about 30 minutes to solve on a 16-core laptop.

Specifically, the process of simulating load adaptation initiated by datacenters is:

- (1) Initial day-ahead dispatch: solve grid DC-OPF with datacenters as fixed loads, which produces grid metrics that guide load adaptation.
- (2) Each datacenter solves its decoupling management LP to propose $\{gridLoad_{i,t}\}$. For PS-GridScale, the grid collects the proposed load of all datacenters and solves a global optimization to determine $\{\alpha_{i,t}\}$ applied to different datacenters.
- (3) Actual dispatch: solve grid DC-OPF with adapted $\{gridLoad_{i,t}\}$. We regard the output as realized grid metrics and report it in evaluation.

With GridCtrl decoupling management, grid operation is simulated by solving grid DC-OPF with solutions of GridCtrl optimization (5).

5.1.4 Carbon Accounting for Datacenters. In the base case where datacenters are fixed loads, each DC’s carbon emissions are calculated by $\sum_t ACI_t \cdot gridLoad_t$, where ACI_t denotes the grid average carbon intensity in the t -th hour. After adding datacenter flexibility with decoupling, we consider two methods to allocate the grid reduction to datacenters:

- **Grid metrics-based (GM):** datacenter carbon emissions are calculated using grid average carbon intensity before and after datacenter flexibility is added.
- **Action-based (Act):** all the grid carbon reduction is allocated to datacenters proportional to their decoupling energy. Formally, on day d , datacenter i is allocated:

$$\Delta gridCarbon_d \cdot \frac{decpen_{i,d}^- + decpen_{i,d}^+}{\sum_{i \in DC} (decpen_{i,d}^- + decpen_{i,d}^+)}$$

The major difference between these two methods is that Act allocates all the grid carbon reduction to datacenters. Such attribution is considered more equitable by many researchers [7, 92] as datacenter flexibility accounts for the grid carbon reduction. Both of the methods would require cooperation with the grid as they need metrics from counterfactual dispatch with no DC flexibility.

5.1.5 Evaluation Metrics. We assess the following metrics from the perspectives of both the grid and datacenters:

- **Grid Dispatch Cost (\$):** the objective of DC-OPF and measures overall grid welfare.

- **Grid Carbon Emissions (kg CO₂)** are calculated by

$$\sum_{t=1}^{24} \sum_f gen_{f,t} * emRate_f$$

where $gen_{f,t}$ is generation from fuel f in the t -th hour and $emRate_f$ is fuel f 's emission rate.

- **Datacenter Carbon Emissions (kg CO₂)**: grid carbon emissions allocated to datacenters. Calculated with grid average carbon intensity (fixed-load) and either GM or Act allocation (flexible). Assuming hyperscalers need to purchase high-quality carbon credits to offset carbon emissions, they can be monetized at \$280/metric ton CO₂ (direct air capture [12]).
- **Datacenter Power Cost (\$)**: the cost of purchasing power from the wholesale power market calculated based on locational prices (LMPs).
- **Decoupling Cost (\$)**: the cost of implementing decoupling. We consider Li-ion battery energy storage (BES), a mature and widely-usable technology, and calculate its total cost of ownership (TCO) based on its static capacity and actual usage needed for decoupling (details in Appendix B). It can be substituted by advancing cheaper storage technologies such as gravity energy storage [90].

The metrics with fixed-load datacenters (current practice) are the baseline when examining the impacts of datacenter flexibility and decoupling. For each metric, we report the average across the day scenarios with weekday/weekend day weighed by 5:2.

5.2 Distributing Decoupling

To study distribution of decoupling, we consider two distribution approaches: EvenDist (equal) and OptDist (grid-optimized), and vary the total decoupling budget. We present results in terms of the normalized total budget (normalized to $\sum_{i \in DC} decpEn_{i,max}^-$ produced by GridCtrl without the budget constraint). Figure 5 shows the distribution results, with the stars representing the values for EvenDist (identical for all DCs), and the whiskers showing the distributions of $decpEn_{i,max}^-$ across DCs for OptDist.

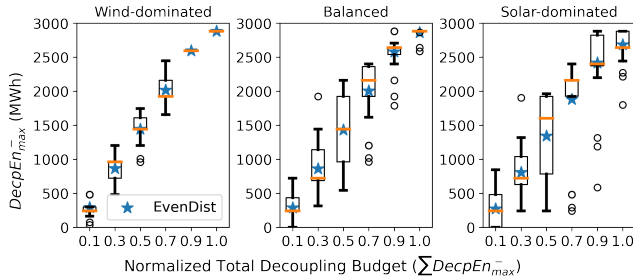


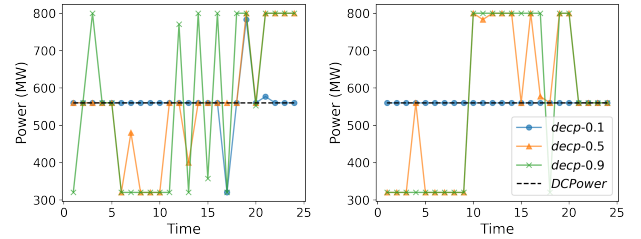
Figure 5: Decoupling Capacity Distributed to Various Datacenters by EvenDist (star) and OptDist (whiskers).

The spread of the whiskers around the star reflects the difference between EvenDist and OptDist. In most cases, an uneven distribution is needed to achieve maximum grid benefits. The distribution

skew increases with more solar generation, because of its temporal correlation (daytimes), leading to transmission congestion. The top-bottom (p95-p5) ratio can be as large as 6–8x, so if datacenters had to pay for their own decoupling, some would have much larger bills. Fundamentally, the skew reflects site heterogeneity due to temporal correlation and transmission congestion, both of which grow as renewable fraction increases.

5.3 Decoupling Impact and Benefits

We illustrate the impact of decoupling on both datacenter and grid dispatch in Figure 6. First, the *DCPower* lines show that stable power is delivered to all datacenters, all the time. Next the “*decp-x*” lines show the impact of various quantities of decoupling capacity. As decoupling capacity is increased, the grid load has more flexibility and evidently changes much more often. These are orchestrated by grid dispatch for benefits such as reduced dispatch cost.



(a) Wind-dominated Grid.

(b) Solar-dominated Grid.

Figure 6: Datacenter Grid Load Variation with Various Decoupling Capacities (OptDist).

Figure 7 shows grid performance versus total decoupling capacity, starting from zero decoupling. As DC grid load flexibility increases (increased decoupling budget), both the grid dispatch cost and carbon emissions are reduced. This is because the decoupling reduces renewable curtailment and load shedding, avoiding dispatch penalties. Carbon emissions are reduced with better renewable absorption. With maximum decoupling budgets, across the three grid types, dispatch cost is reduced by 20–38%, and carbon emissions are reduced by 10–17%. There are diminishing returns—70% of the maximum budget captures 98–100% of the grid benefits across grid types, implying that significant cost savings are possible with little loss in grid benefits. Decoupling benefits are higher in the balanced and solar-dominated grids where solar generation produces a greater need for DC flexibility.

The difference in grid performance with OptDist and EvenDist is consistent with the distribution difference. In the wind-dominated grid where the skew of distribution is smaller, OptDist improves grid performance less (Figure 8a); in the other two grids and with medium total budgets (0.3–0.7), the advantage of OptDist gets more evident. Figure 8b shows additional grid carbon reduction of 0.7% (2,000 metric ton CO₂/day). For the following studies, the results are based on OptDist if not specified.

5.4 Managing Decoupling Cooperatively

We know that GridCtrl, centralized grid control of decoupling, will produce the maximum grid benefits. However, if decoupling is

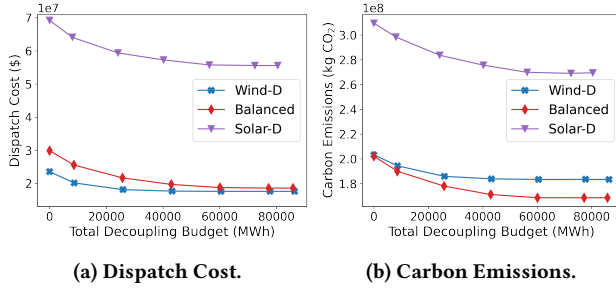


Figure 7: Grid Performance vs. Total Decoupling Capacity. (OptDist distribution)

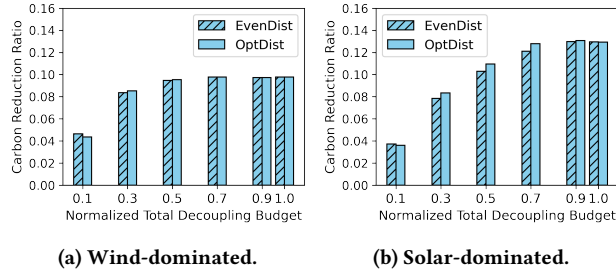


Figure 8: Grid Carbon Reduction with Even and Optimized Distribution of Decoupling.

owned or paid for by datacenters, then the operators may prefer to preserve some management autonomy. So we consider two schemes that retain datacenter autonomy, PlanShare and PS-GridScale, which allow datacenters to control their decoupling resources for local benefits as well. The key question is: how do these schemes perform—for datacenter and grid?

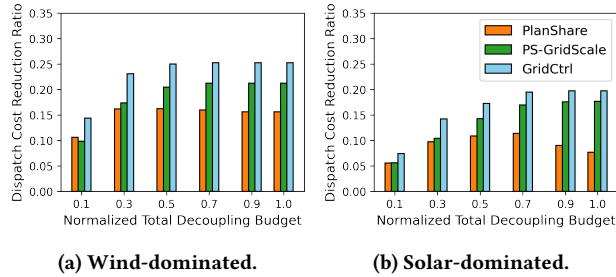


Figure 9: Grid Dispatch Cost Reduction with Various Decoupling Management Approaches.

In Figure 9, we consider grid dispatch cost. And in Figure 10 we consider grid carbon reduction. For both metrics (w/ OptDist distribution), PlanShare (orange) fails to manage decoupling effectively—the benefits are generally lower than the two other schemes, and can even decrease as the total budget increases. Although still falling short of GridCtrl performance, PS-GridScale (green) consistently outperforms PlanShare, achieving 84/90% of the dispatch cost reduction and 84/89% of the carbon reduction with GridCtrl in the

wind/solar-dominated grid. These are up to 1.6x better dispatch cost reduction and 1.4x better carbon reduction vs. PlanShare. These two approaches also see diminishing returns in decoupling capacity—most of the benefits are captured with 70% of the maximum total budget.

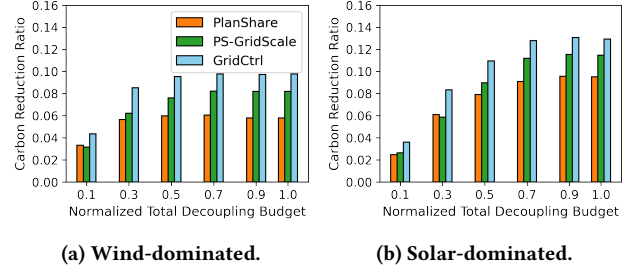


Figure 10: Grid Carbon Reduction with Various Decoupling Management Approaches.

To understand the performance difference between PlanShare and PS-GridScale, we examine their decoupling behaviors. Figure 11 shows the time series for grid load of 10 sampled datacenters. With PlanShare, correlated local prices (LMP) cause highly correlated grid load profiles (Figure 11a), collectively resulting in large load variation (overshifting) that disrupts grid operation. The disruption is reflected as increased gas generation in the wind-dominated grid and load shedding in the solar-dominated grid. With PS-GridScale, although the load profiles initially proposed are also correlated on LMPs, the introduction of grid-decided scale factors breaks the correlation and thus mitigates overshifting (Figure 11b). The overall performance advantage demonstrates the necessity of cooperation between large flexible loads and the grid.

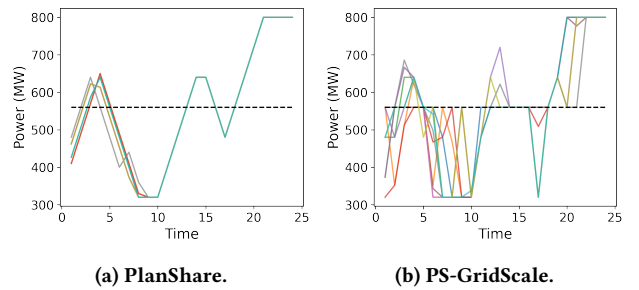


Figure 11: Datacenter Grid Load Examples with PlanShare and PS-GridScale Management.

We also revisit the impact of decoupling distribution with PlanShare and PS-GridScale. In the wind-dominated grid, the performance impact is minor, so in Figure 12 we compare grid dispatch cost in the solar-dominated grid. PlanShare clearly benefits from OptDist, the better distribution, while PS-GridScale is robust against distribution variation.

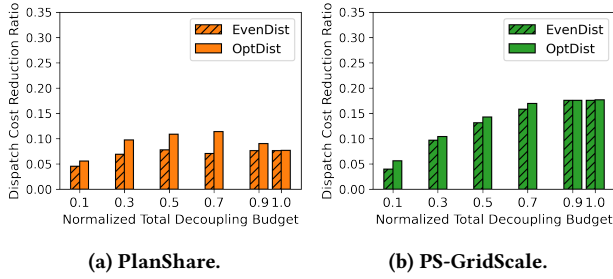


Figure 12: Comparing Dispatch Cost for Even (EvenDist) and Optimized Decoupling Distribution (OptDist) with PlanShare and PS-GridScale. Solar-dominated Grid.

5.5 Economic Benefits of Decoupling

To figure out whether decoupling is economically viable, we assess the economic benefits and compare them with decoupling cost for datacenters and the whole grid.

5.5.1 Datacenter Economic Benefits from Decoupling. With decoupling, datacenters get two economic benefits: (1) reduced power cost and (2) reduced carbon emissions. We compare the average benefits normalized to average decoupling costs (70% total budget, OptDist distribution) in Figure 13. Power cost savings alone are significant. With PS-GridScale (Figure 13a, “PC-R”), the ratio of power cost reduction to decoupling cost is 0.4–3, which is especially high in the solar-dominated grid where the power price is higher. With the avoided cost of carbon credits considered, grid metrics-based (GM) allocation brings the total cost savings to 0.8–3.6x of decoupling cost. Action-based (Act) allocation allocates more grid carbon reduction to datacenters, and directly the average economic benefits to datacenters exceed their decoupling cost. GridCtrl can improve the benefit-cost ratio in the balanced grid by further lowering datacenter power cost.

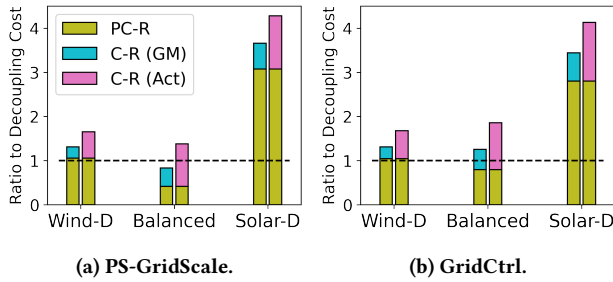


Figure 13: Average Datacenter Benefits vs. Decoupling Cost. (PC-R: power cost reduction, C-R: carbon reduction)

We consider individual datacenters’ carbon reduction benefits, using the two different reduction allocation methods, GM and Act, in Figure 14. With GM allocation, datacenters each see a carbon reduction of about 10%, mostly due to decreased in grid carbon intensity. Act allocation attributes all grid carbon reduction to datacenters, proportional to local decoupling capacity. In this case, most see a >2x reduction, with outliers low due to low local decoupling capacity.

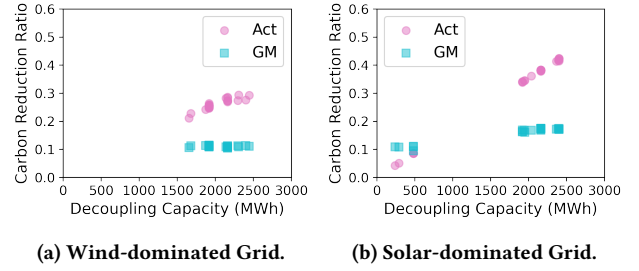


Figure 14: Carbon Reduction (GridCtrl) with Grid Metrics- and Action-based Allocation at Individual Datacenters.

Datacenters may have different owners⁵, so differences amongst them are important. Figure 15 compares the ratio of economic benefits (power cost savings and carbon reduction) to decoupling cost at individual datacenters. In the wind-dominated grid, all datacenters see a net gain (ratio>1) from load decoupling with either management approach. In the balanced grid, most datacenters experience an even higher benefit-cost ratio, but there is a wide range. In the solar-dominated grid, the average ratio is even higher, but there is an even wider range. In both balanced and solar-dominated grids, some datacenters see very negative returns, and thus are unlikely to deploy decoupling of their own initiative. Overall, these results suggest that voluntary programs are unlikely to produce optimal (or even good) decoupling distributions, and some differential incentive programs will be needed to achieve OptDist-like distributions.

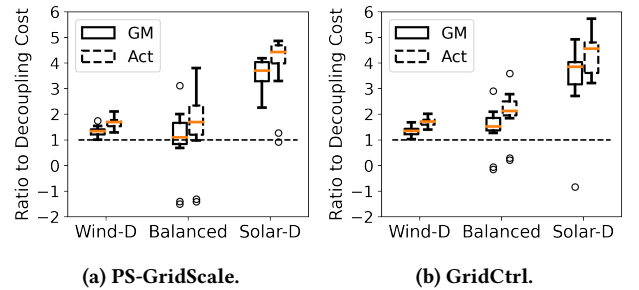


Figure 15: Ratio of Economic Benefits to Decoupling Cost at Individual Datacenters.

5.5.2 Grid Benefits and Costs. An alternative to decoupling is to add more renewable generation, thereby advancing grid decarbonization. In fact, some studies suggest that overprovisioning renewables is the cheapest way to reduce carbon emissions [34, 71]. We compare the effectiveness of additional renewable generation against decoupling.

Figure 16 compares the annualized generation cost⁶ with decoupling (70% total budget, OptDist distribution). With PS-GridScale, the costs are the same in wind-dominated grid, but decoupling is 40% cheaper in balanced grid and enables reduction unachievable with only more solar generation. GridCtrl slightly increases

⁵And OptDist produces different decoupling investment needs.

⁶Estimated with leveled cost of energy (LCOE) of \$50/MWh for wind and \$60/MWh for solar (2023 midpoints [50]).

the decoupling cost due to more battery usage, but as the carbon benefits improve more, its decoupling cost is 8% lower in the wind-dominated grid and 46% lower in the balanced grid.

Overall, for grids with 60% renewables, implementing datacenter load decoupling (flexibility) is more economic for decarbonization than adding more renewables.

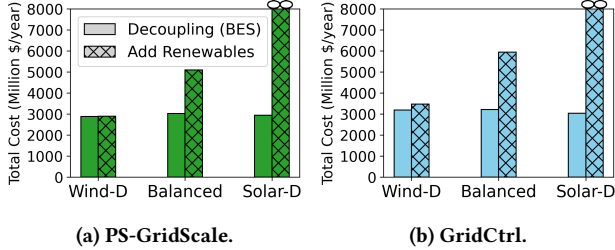


Figure 16: Comparison of Grid Carbon Reduction Cost (∞ : carbon reduction unachievable with adding renewables).

6 Related Work

6.1 Datacenter Flexibility

Load Decoupling. Previous work that controls energy resources co-located with datacenters for load adaptation shares the idea of load decoupling. Some explore reusing widely-equipped UPS energy storage for grid services such as frequency regulation and peak shaving, reducing power cost and aiding the grid [68, 78, 88, 99], but the energy capacity is too low (typically ≈ 10 minutes) to decouple datacenter power and grid load for hours—scenarios in this paper. [2, 74] assessed how much renewable power purchase agreement (PPA) and energy storage would be needed for various renewable matching goals (e.g. 100% hourly matching) at individual DCs. We do extensive DC-grid coupled simulations that capture complex dynamics to explore decoupling needs and examine both DC and grid benefits.

Workload Flexibility. Although we assume datacenter power capacity needs to be constant as is typical with cloud computing and important for high DC capital efficiency, some workloads are time- and space-shiftable (e.g. machine learning model training and batch processing jobs). Shifting enables flexible DC capacity without QoS loss. These workloads can be significant in some commercial cloud workloads (e.g. 25% of CPU core usage at Microsoft) [2, 70, 72]. Many researchers have exploited scheduling or scaling these workloads to match datacenter grid load to low-priced/low-carbon generation for cost or carbon reduction [2, 30, 36, 55, 57, 58, 72, 84, 93, 95]. However, all of these studies assume space capacity, an undesirable situation from an economic or business point-of-view. Workload flexibility is complementary to load decoupling, and correct exploitation can reduce both energy deficit and decoupling cost.

6.2 Coordinating Adaptive Datacenters and Grid

Our exploration of managing decoupling for efficiency is related to coordinating DC flexibility and grid dispatch. [28, 53, 64] show the benefits of exploiting load flexibility for grid resource adequacy or datacenter growth but do not address the coordination problem.

[32, 54, 56] show that selfish control (local optimization) can harm the grid and datacenters themselves when DCs consume 10–20% of grid energy (or more). With the growth of ML/AI, more power grids are seeing DC load fractions that exceed 15% [40, 66] (or even 30% [24, 33]), and thus coordination is becoming increasingly important. Several studies of coordination explore information sharing and various DC-grid interaction designs, studying load adaptation with grid simulation [32, 52, 56, 61, 94, 101, 102]. Specifically, we show that 1-way info sharing represented by PlanShare [52] can be ineffective as DCs continue to grow. Addressing this, we design PS-GridScale that empowers the grid to participate in load shaping, enabling it to outperform PlanShare significantly.

7 Summary and Future Work

With the rapid growth of datacenter load, DC load flexibility is critical in reducing datacenter carbon emissions. Specifically, it can resolve the conflict between DC growth and grid decarbonization. To protect DC compute efficiency and QoS, such flexibility requires decoupling compute capacity and grid load. With the goal of achieving high benefit at low cost, we study how to best distribute and manage decoupling. Our evaluation shows that $>98\%$ of maximum benefits can be achieved with 70% of the maximum decoupling need. Cooperation between the grid and DCs is necessary for efficient decoupling management: 2-way sharing and control represented by PS-GridScale enables 1.4x grid carbon reduction vs. 1-way information sharing and achieves 84–90% of the optimal benefits. Finally, we show that decoupling can be economically viable, the economic benefits to datacenters, on average, exceed their decoupling costs. However, significant skew across different datacenters—some may not benefit at all or enough to defray their costs—means a significant grid intervention may be required to implement decoupling.

Our framework of datacenter load decoupling opens up many research opportunities. When datacenters propose load profiles, can there be better grid metrics that guide effective shaping and avoid oversubscribing the opportunities? Allocation methods that fairly allocate the grid benefits and are invulnerable to malicious load profiles are important for formulating grid programs. Discussion beyond technical problems such as who should pay for decoupling is also important.

Acknowledgments

This work is supported in part by NSF Grants OAC-2442555, CNS-2325956, and the VMware University Research Fund. We also thank the Large-scale Sustainable Systems Group members for their support of this work!

References

- [1] 2019. Clicking Clean Virginia: The Dirty Energy Powering Data Center Alley. <https://www.greenpeace.org/usa/reports/click-clean-virginia/>.
- [2] Bilge Acun, Benjamin Lee, Fiodar Kazhamiaka, Kiwan Maeng, Udit Gupta, Manoj Chakkaravarthy, David Brooks, and Carole-Jean Wu. 2023. Carbon explorer: A holistic framework for designing carbon aware datacenters. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*. 118–132.
- [3] U.S. Energy Information Administration. 2024. Average Power Plant Operating Expenses for Major U.S. Investor-Owned Electric Utilities, 2013 through 2023. https://www.eia.gov/electricity/annual/html/epa_08_04.html
- [4] Ranganathan Barroso, Holze. 2018. *The Datacenter As A Computer: Designing Warehouse-Scale Machines, Third Edition* (third ed.). Morgan-Claypool.
- [5] Oliver Blake. 2025. Microsoft's Q4 2025 Earnings: Navigating AI's Short-Term Costs for Long-Term Dominance. <https://www.ainvest.com/news/microsoft-q4-2025-earnings-navigating-ai-short-term-costs-long-term-dominance-2507/>
- [6] New York State Energy Planning Board. 2015. The Energy to Lead: 2015 New York State Energy Plan. <http://energyplan.ny.gov/Plans/2015.aspx>.
- [7] Francesca Bona, Nicolas Gast, Jean-Yves Le Boudec, Pierre Pinson, and Dan-Cristian Tomozei. 2016. Attribution mechanisms for ancillary service costs induced by variability in power delivery. *IEEE Transactions on Power Systems* 32, 3 (2016), 1891–1901.
- [8] Bruno Burger and Claudia Hanisch. 2025. German Net Power Generation in 2024: Electricity Mix Cleaner than Ever. <https://www.ise.fraunhofer.de/en/press-media/press-releases/2025-public-electricity-generation-2024-renewable-energies-cover-more-than-60-percent-of-german-electricity-consumption-for-the-first-time.html>
- [9] Andrew A Chien and Liuzixuan Lin. 2024. As Grids Reach 100% Renewable at Peak, Growing Curtailment of 8 Gigawatts Looms as a Challenge to Decarbonization. *ACM SIGENERGY Energy Informatics Review* 4, 1 (2024), 3–10.
- [10] 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.
- [11] Andrew A Chien, Chaojie Zhang, Liuzixuan Lin, and Varsha Rao. 2022. Beyond PUE: Flexible datacenters empowering the cloud to decarbonize. *USENIX Hot Carbon* (2022).
- [12] Heather Clancy. 2024. How Google negotiated the lowest price yet for a direct air carbon capture contract. <https://trellis.net/article/how-google-negotiated-the-lowest-price-yet-for-a-direct-air-carbon-capture-contract/>
- [13] Wesley Cole and Akash Karmakar. 2023. *Cost projections for utility-scale battery storage: 2023 update*. Technical Report. National Renewable Energy Lab.(NREL), Golden, CO (United States).
- [14] California Public Utilities Commission. 2021. CPUC Drives California Toward a Clean Energy Future. <https://www.cpuc.ca.gov/news-and-updates/all-news/cpuc-drives-california-toward-a-clean-energy-future-2024>
- [15] A committee of faculty and staff. 2021. *The Timeline and Events of the February 2021 Texas Electric Grid Blackouts*. Technical Report. The University of Texas at Austin.
- [16] McKinsey & Company. 2024. AI power: Expanding data center capacity to meet growing demand. <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/ai-power-expanding-data-center-capacity-to-meet-growing-demand/>
- [17] Michael J. Coren. 2016. Germany had so much Renewable Energy on Sunday that it had to pay people to use Electricity. <http://qz.com/680661/germany-had-so-much-renewable-energy-on-sunday-that-it-had-to-pay-people-to-use-electricity/>
- [18] Microsoft Corporation. 2025. Microsoft Environmental Sustainability Report 2025. <https://www.microsoft.com/en-us/corporate-responsibility/sustainability/report/>
- [19] 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.
- [20] European Commission. Climate Action DG. 2019. *Going Climate-neutral by 2050: A Strategic Long-term Vision for a Prosperous, Modern, Competitive and Climate-neutral EU Economy*. Publications Office of the European Union.
- [21] 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.
- [22] 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/>
- [23] Iain Dunning, Joey Huchette, and Miles Lubin. 2017. JuMP: A modeling language for mathematical optimization. *SIAM review* 59, 2 (2017), 295–320.
- [24] Virginia Electric and Power Company. 2023. 2023 Integrated Resource Plan. <https://www.dominionenergy.com/about/our-company/irp#>.
- [25] Department for Energy Security, Net Zero, and Energy & Industrial Strategy Department for Business. 2021. Net Zero Strategy: Build Back Greener.
- [26] Policy Research Center for Environment, Ministry of Ecology Economy (PRCEE), Environment, and Environmental Defense Fund Beijing Representative Office. 2024. *China's Policies and Actions on Carbon Peaking and Carbon Neutrality*. Technical Report. Policy Research Center for Environment and Economy (PRCEE).
- [27] Christopher Gannatti. 2024. This Is What AI Commitment Looks Like: \$392 Billion and Rising. <https://www.wisdomtree.com/investments/blog/2025/05/21/this-is-what-ai-commitment-looks-like-392-billion-and-rising>
- [28] Brian F Gerke, Sarah Josephine Smith, Samanvitha Murthy, Sunhee Baik, Shreya Agarwal, Peter Alstone, Aditya Khandekar, Cong Zhang, Richard E Brown, Jingjing Liu, and Mary Ann Piette. 2024. The California Demand Response Potential Study, Phase 4: Report on Shed and Shift Resources Through 2050.
- [29] Wedan Emmanuel Gribga, Anne Blavette, and Anne-Cécile Orgerie. 2023. Renewable energy in data centers: the dilemma of electrical grid dependency and autonomy costs. *IEEE Transactions on Sustainable Computing* 9, 3 (2023), 315–328.
- [30] Ñigo Goiri, William Katsak, Kien Le, Thu D Nguyen, and Ricardo Bianchini. 2013. Parosol and greenswitch: Managing datacenters powered by renewable energy. In *ACM SIGARCH Computer Architecture News*. ACM, 51–64.
- [31] Google. 2025. Google 2025 Environmental Report. <https://sustainability.google/google-2025-environmental-report/>
- [32] Joseph Gorka, Noah Rhodes, and Line Roald. 2025. Electricityemissions. jl: A framework for the comparison of carbon intensity signals. In *Proceedings of the 16th ACM International Conference on Future and Sustainable Energy Systems*. 19–30.
- [33] EirGrid Group. 2024. Generation Capacity Statement 2023–2032. <https://cms.eirgrid.ie/sites/default/files/publications/19035-EirGrid-Generation-Capacity-Statement-Combined-2023-V5-Jan-2024.pdf>.
- [34] Rahul Gupta and Fabrizio Sossan. 2023. Optimal sizing and siting of energy storage systems considering curtailable photovoltaic generation in power distribution networks. *Applied Energy* 339 (2023), 120955.
- [35] Gurobi Optimization, LLC. 2023. Gurobi Optimizer Reference Manual. <https://www.gurobi.com>
- [36] Walid A Hanafy, Qianlin Liang, Noman Bashir, David Irwin, and Prashant Shenoy. 2023. Carbonscaler: Leveraging cloud workload elasticity for optimizing carbon-efficiency. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* 7, 3 (2023), 1–28.
- [37] M Huneault and FD Galiana. 1991. A survey of the optimal power flow literature. *IEEE transactions on Power Systems* 6, 2 (1991), 762–770.
- [38] International Energy Agency (IEA). 2024. Electricity 2024. <https://www.iea.org/reports/electricity-2024>
- [39] International Energy Agency (IEA). 2024. Renewables 2024. <https://www.iea.org/reports/renewables-2024/electricity>
- [40] Electric Power Research Institute. 2024. Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption. <https://www.epri.com/research/products/3002028905>
- [41] Electric Power Research Institute. 2025. DCFlex Initiative. <https://dcflex.sf.epri.com/>
- [42] Ponemon Institute. 2016. Cost of Data Center Outages. https://www.vertiv.com/globalassets/documents/reports/2016-cost-of-data-center-outages-11-11_51190_1.pdf
- [43] California ISO. [n. d.]. ISO 2022-2023 Transmission Plan, Appendix F: Detailed Policy Assessment. https://stakeholdercenter.caiso.com/InitiativeDocuments/Appendix-F-Board-Approved_2022-2023_Transmission-Plan.pdf
- [44] California ISO. 2023. Special Report on Battery Storage. <https://www.caiso.com/documents/2022-special-report-on-battery-storage-jul-7-2023.pdf>
- [45] California ISO. 2024. Monthly Renewables Performance Report, Dec. 2024. <https://www.caiso.com/documents/monthly-renewables-performance-report-december-2024.html>
- [46] Rabih A Jabr, Izudin Džafić, and Bikash C Pal. 2014. Robust optimization of storage investment on transmission networks. *IEEE Transactions on Power Systems* 30, 1 (2014), 531–539.
- [47] 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/>
- [48] Peter Judge. 2022. Singapore lifts data center moratorium - but sets conditions. <https://www.datacenterdynamics.com/en/news/singapore-lifts-data-center-moratorium-but-sets-conditions/>
- [49] Peter Judge. 2023. Microsoft granted permission to run its Dublin data center on gas. <https://www.datacenterdynamics.com/en/news/microsoft-granted-permission-to-run-its-dublin-data-center-on-gas/>
- [50] Lazard. 2025. Lazard's 2025 Levelized Cost Of Energy+. <https://www.lazard.com/media/uoounhon4/lazards-lcoepplus-june-2025.pdf>
- [51] 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.
- [52] Liuzixuan Lin and Andrew A Chien. 2023. Adapting Datacenter Capacity for Greener Datacenters and Grid. In *Proceedings of the 14th ACM International*

- Conference on Future Energy Systems. 200–213.
- [53] Liuzixuan Lin, Rajini Wijayawardana, Varsha Rao, Hai Nguyen, Emmanuel Wedan GNIBGA, and Andrew A Chien. 2024. Exploding AI Power Use: an Opportunity to Rethink Grid Planning and Management. In *Proceedings of the 15th ACM International Conference on Future and Sustainable Energy Systems*. 434–441.
 - [54] 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.
 - [55] 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.
 - [56] 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.
 - [57] 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.
 - [58] 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.
 - [59] Kari Lydersen. 2024. Data centers offer energy peril and promise, with the Midwest increasingly in the crosshairs. <https://www.minnpost.com/other-nonprofit-media/2024/06/microsoft-data-centers-energy-peril-promise-midwest-energy-water-concerns/>
 - [60] Eric Masanet, Arman Shehabi, Nuo Lei, Sarah Smith, and Jonathan Koomey. 2020. Recalibrating global data center energy-use estimates. *Science* 367, 6481 (2020), 984–986. [arXiv:https://science.sciencemag.org/content/367/6481/984.full.pdf](https://arxiv.org/abs/https://science.sciencemag.org/content/367/6481/984.full.pdf) doi:10.1126/science.aba3758
 - [61] 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).
 - [62] Sebastian Moss. 2025. Meta to invest "hundreds of billions of dollars into compute to build superintelligence," with several multi-GW data center clusters. <https://www.datacenterdynamics.com/en/news/meta-to-invest-hundreds-of-billions-of-dollars-into-compute-to-build-superintelligence-with-several-multi-gw-data-center-clusters/>
 - [63] Samuel Newell, Michael Hagerty, Johannes Pfeifenberger, Bin Zhou, Travis Carless, Rohan Janakiraman, Sang Gang, Patrick Daou, and Joshua Junge. 2022. PJM CONE 2026/2027 Report. (2022).
 - [64] Tyler Norris, Timothy Profeta, Dalia Patino-Echeverri, and Adam Cowie-Haskell. 2025. *Rethinking Load Growth: Assessing the Potential for Integration of Large Flexible Loads in US Power Systems*. Technical Report. Nicholas Institute for Energy, Environment & Sustainability, Duke University.
 - [65] Electric Reliability Council of Texas. [n. d.]. Large Flexible Load Task Force. <https://www.ercot.com/committees/tac/lfltf>.
 - [66] Electric Reliability Council of Texas. 2025. Load Forecast. <https://www.ercot.com/gridinfo/load/forecast>
 - [67] Isabella O'Malley. 2025. China built out record amount of wind and solar power in 2024. <https://apnews.com/article/wind-solar-energy-china-climate-carbon-emissions-b337503abfacd9b7829fd7bbcd507e9>
 - [68] Janne Paananen and Ehsan Nasr. 2021. Grid-interactive data centers: enabling decarbonization and system stability. <https://www.eaton.com/content/dam/eaton/markets/data-center/eaton-microsoft-grid-interactive-whitepaper-wp153031en.pdf>
 - [69] 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.
 - [70] Anjaly Parayil, Jue Zhang, Xiaoting Qin, Íñigo Goiri, Lexiang Huang, Timothy Zhu, and Chetan Bansal. 2024. Towards Cloud Efficiency with Large-scale Workload Characterization. *arXiv preprint arXiv:2405.07250* (2024).
 - [71] Morgan Putnam and Marc Perez. 2018. Solar Potential Analysis Report. Prepared for Minnesota Department of Commerce and the Minnesota Solar Pathways Project (curtailment more cost-effective than storage), Available from <http://mnsolarpathways.org/>.
 - [72] Ana Radovanović, Ross Koningstein, Ian Schneider, Bokan Chen, Alexandre Duarte, Bin Yue Xiao, Diyu Xia, Maya Haridasan, Patrick Hung, Nick Care, et al. 2022. Carbon-aware computing for datacenters. *IEEE Transactions on Power Systems* 38, 2 (2022), 1270–1280.
 - [73] Benjamin Reidys, Pantea Zardoshti, Íñigo Goiri, Celine Irvine, Daniel S Berger, Haoran Ma, Kapil Arya, Eli Cortez, Taylor Stark, Eugene Bak, et al. 2025. Coach: Exploiting temporal patterns for all-resource oversubscription in cloud platforms. In *Proceedings of the 30th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 1*. 164–181.
 - [74] Chuangang Ren, Di Wang, Bhuvan Uргаonkar, and Anand Sivasubramaniam. 2012. Carbon-aware energy capacity planning for datacenters. In *2012 IEEE 20th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems*. IEEE, 391–400.
 - [75] Goldman Sachs Research. 2024. AI/data centers' global power surge and the Sustainability impact. <https://www.goldmansachs.com/images/migrated/insights/pages/gs-research/gs-sustain-generational-growth-ai-data-centers-global-power-surge-and-the-sustainability-impact/sustain-data-center-redaction.pdf>
 - [76] Ian Schneider and Taylor Mattia. 2024. Carbon accounting in the Cloud: a methodology for allocating emissions across data center users. *arXiv preprint arXiv:2406.09645* (2024).
 - [77] Arman Shehabi, Sarah Josephine Smith, Alex Hubbard, Alexander Newkirk, Nuo Lei, Md AbuBakar Siddik, Billie Holecck, Jonathan G Koomey, Eric R Masanet, and Dale A Sartor. 2024. *2024 United States Data Center Energy Usage Report*. Technical Report. Lawrence Berkeley National Laboratory.
 - [78] 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.
 - [79] João da Silva, Natalie Sherman, and Imran Rahman-Jones. 2025. Tech giants are putting \$500bn into 'Stargate' to build up AI in US. <https://www.bbc.com/news/articles/cy4m84d2xz2o>
 - [80] Patrick Sisson. 2024. A.I. Frenzy Complicates Efforts to Keep Power-Hungry Data Sites Green. <https://www.nytimes.com/2024/02/29/business/artificial-intelligence-data-centers-green-power.html>
 - [81] Zachary Skidmore. 2025. xAI deploys 168 Tesla Megapacks to power its Colossus supercomputer in Memphis. <https://www.datacenterdynamics.com/en/news/xai-deploys-168-tesla-megapacks-to-power-its-colossus-supercomputer-in-memphis/>
 - [82] Abel Souza, Noman Bashir, Jorge Murillo, Walid Hanafy, Qianlin Liang, David Irwin, and Prashant Shenoy. 2023. Ecovisor: A virtual energy system for carbon-efficient applications. In *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*. 252–265.
 - [83] Jim Steel. 2015. The What, When And How of Texas Electricity Prices Going Negative. <https://cleantechnica.com/2015/10/01/texas-electricity-prices-going-negative/>
 - [84] Thanathorn Sukprasert, Abel Souza, Noman Bashir, David Irwin, and Prashant Shenoy. 2024. On the Limitations of Carbon-Aware Temporal and Spatial Workload Shifting in the Cloud. In *Proceedings of the Nineteenth European Conference on Computer Systems*. 924–941.
 - [85] Dan Swinhoe. 2024. Amazon invests in nuclear SMR company, signs several SMR deals to power data centers. <https://www.datacenterdynamics.com/en/news/amazon-signs-deals-to-invest-in-nuclear-smrs-to-power-data-centers/>
 - [86] Dan Swinhoe. 2024. Utilities Oncor, Xcel, and Ameren see increased data center capacity demands over Q2. <https://www.datacenterdynamics.com/en/news/utilities-oncor-xcel-and-ameren-see-increased-data-center-capacity-demands-over-q2/>
 - [87] 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. doi:10.1145/3342195.3387517
 - [88] Rahul Uргаonkar, Bhuvan Uргаonkar, 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.
 - [89] Deepthi Vaidhyanathan, Kumaraguru Prabakar, Gregory Martin, Anand Ramesh, Ben Wheeler, Christopher Coco, Jimmy Clidaras, Matthieu Monsch, Sangsun Kim, Saurav Talukdar, et al. 2025. *Vulcan Test Platform: Demonstrating the Data Center as a Flexible Grid Asset*. Technical Report. National Renewable Energy Laboratory.
 - [90] Vilayanur Viswanathan, Kendall Mongird, Ryan Franks, Xiaolin Li, Vincent Sprenkle, and R Baxter. 2022. 2022 grid energy storage technology cost and performance assessment. *Energy* 2022 (2022).
 - [91] Di Wang, Chuangang Ren, Anand Sivasubramaniam, Bhuvan Uргаonkar, and Hosam Fathy. 2012. Energy storage in datacenters: what, where, and how much?. In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*. 187–198.
 - [92] Lucien Werner, Adam Wierman, and Steven H Low. 2021. Pricing flexibility of shiftable demand in electricity markets. In *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. 1–14.
 - [93] Philipp Wiesner, Ilja Behnke, Dominik Scheinert, Kordian Gontarska, and Lauritz Thamsen. 2021. Let's wait awhile: how temporal workload shifting can

reduce carbon emissions in the cloud. In *Proceedings of the 22nd International Middleware Conference*. 260–272.

- [94] Zhaoyuan Wu, Lin Chen, Jianxiao Wang, Ming Zhou, Gengyin Li, and Qing Xia. 2023. Incentivizing the spatiotemporal flexibility of data centers toward power system coordination. *IEEE Transactions on Network Science and Engineering* (2023).
- [95] Jiali Xing, Bilge Acun, Aditya Sundararajan, David Brooks, Manoj Chakkaravarthy, Nikky Avila, Carole-Jean Wu, and Benjamin C Lee. 2023. Carbon Responder: Coordinating Demand Response for the Datacenter Fleet. *arXiv preprint arXiv:2311.08589* (2023).
- [96] Peng Xiong and Chanan Singh. 2015. Optimal planning of storage in power systems integrated with wind power generation. *IEEE Transactions on Sustainable Energy* 7, 1 (2015), 232–240.
- [97] Qi Ye, Lu Jiaqi, and Zhu Mengye. 2018. Wind curtailment in China and lessons from the United States. *Brookings China's Energy in Transition Series* (2018).
- [98] Haoyu Yuan, Reetam Sen Biswas, Jin Tan, and Yingchen Zhang. 2020. Developing a reduced 240-bus WECC dynamic model for frequency response study of high renewable integration. In *2020 IEEE/PES transmission and distribution conference and exposition (T&D)*. IEEE, 1–5.
- [99] Baosen Zhang and Ehsan Nasr Senior. 2021. Distributed redundant integration of data center battery storage with the grid for regulation services. In *2021 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 1–5.
- [100] Chaojie Zhang and Andrew A Chien. 2021. Scheduling Challenges for Variable Capacity Resources. In *Workshop on Job Scheduling for Parallel Processing (JSSPP)*.
- [101] 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.
- [102] 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.

A Formulation and Settings of Direct-current Optimal Power Flow (DC-OPF)

Here we present the complete formulation of DC-OPF problem used for grid simulation, starting from the notations:

Table 2: DC-OPF Notations: Sets

Notation	Description	Notation	Description
$\mathcal{T}$	Time periods	$\mathcal{N}$	Buses
$\mathcal{G} (\mathcal{G}_n)$	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
$\mathcal{DC}$	Datacenters (at bus n)	$\mathcal{ND}$	Non-DC loads (at bus n)
$(\mathcal{DC}_n)$		$(\mathcal{ND}_n)$	
$\mathcal{W} (\mathcal{W}_n)$	Wind farms (at bus n)	$\mathcal{S} (\mathcal{S}_n)$	Solar farms (at bus n)
$\mathcal{R} (\mathcal{R}_n)$	Other renewable generators (at bus n)		

Table 3: DC-OPF Notations: Decision Variables

Notation	Description	Notation	Description
$p_{i,t}$	Generation of generator i at time t	$f_{l,t}$	Power flow of line l at time t
$d_{i,t}^{nd}$	Load shedding at non-DC load i at time t	$d_{i,t}^{dc}$	Load shedding at datacenter i at time t
$m_{i,t}$	Curtailment at import i at time t	$w_{i,t}$	Curtailment at wind farm i at time t
$s_{i,t}$	Curtailment at solar farm i at time t	$r_{i,t}$	Curtailment at other renewable i at time t
$\theta_{n,t}$	Phase angle at bus n at time t		

Table 4: DC-OPF Notations: Parameters

Notation	Description	Notation	Description
B_l	Susceptance of transmission line l	C_i	Generation cost of generator i
C_i^{nd}	Load-shedding penalty at non-DC load i	C_i^{dc}	Load-shedding penalty at datacenter i
C_i^w	Curtailment penalty at wind farm i	C_i^s	Curtailment penalty at solar farm i
C_i^m	Curtailment penalty at import point i	C_i^r	Curtailment penalty at other renewable i
$D_{i,t}$	Power demand of load i at time t	F_l^{max}	Maximum power flow of transmission line l
$M_{i,t}$	Power from import i at time t	P_i^{max}	Maximum power output of generator i
RD_i	Ramp-down limit of generator i	RU_i	Ramp-up limit of generator i
$W_{i,t}$	Generation of wind farm i at time t	$S_{i,t}$	Generation of solar farm i at time t
$R_{i,t}$	Generation of other renewable i 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		

The generation cost of thermal generators is 6/31/22 \$/MWh for nuclear/coal/gas generators [3]. Penalties of load shedding and curtailment are listed in Table 5, which are reflected in the locational marginal price (LMP) when load shedding or renewable/import curtailment happens. The other parameters are set according to load/generation profiles or static infrastructure attributes.

Table 5: DC-OPF Parameter Settings

	Value		Value		Value
C_i^{nd}	\$1000/MWh	C_i^{dc}	\$1000/MWh	C_i^w	\$100/MWh
C_i^s	\$100/MWh	C_i^m	\$500/MWh	C_i^r	\$1000/MWh

The optimization objective is to minimize the daily dispatch cost consisting of generation cost, load shedding penalties, and import/renewable generation curtailment penalties:

$$\min \sum_{t \in \mathcal{T}} \left(\sum_{i \in \mathcal{G}} C_i p_{i,t} + \sum_{i \in \mathcal{ND}} C_i^{nd} d_{i,t}^{nd} + \sum_{i \in \mathcal{DC}} C_i^{dc} d_{i,t}^{dc} + \sum_{i \in \mathcal{I}} C_i^m m_{i,t} + \sum_{i \in \mathcal{W}} C_i^w w_{i,t} + \sum_{i \in \mathcal{S}} C_i^s s_{i,t} + \sum_{i \in \mathcal{R}} C_i^r r_{i,t} \right) \quad (6a)$$

subject to typical constraints including balancing at each node (6b), transmission (6c–6e), generator capacity (6f) and ramping (6g), and

shedding/curtailment limits (6h–6m):

$$\begin{aligned}
\text{s.t. } & \sum_{l \in \mathcal{L}_n^+} f_{l,t} - \sum_{l \in \mathcal{L}_n^-} f_{l,t} + \sum_{i \in \mathcal{G}_n} p_{i,t} + \sum_{i \in \mathcal{I}_n} (M_{i,t} - m_{i,t}) \\
& + \sum_{i \in \mathcal{W}_n} (W_{i,t} - w_{i,t}) + \sum_{i \in \mathcal{S}_n} (S_{i,t} - s_{i,t}) + \sum_{i \in \mathcal{R}_n} (R_{i,t} - r_{i,t}) \\
& = \sum_{i \in \mathcal{ND}_n} (D_{i,t} - d_{i,t}^{nd}) + \sum_{i \in \mathcal{DC}_n} (\text{gridLoad}_{i,t} - d_{i,t}^{dc}), \\
& \forall n \in \mathcal{N}, t \in \mathcal{T}, \tag{6b} \\
& f_{l,t} = B_l(\theta_{n,t} - \theta_{m,t}), \quad \forall l = (m, n) \in \mathcal{L}, t \in \mathcal{T}, \tag{6c} \\
& -F_l^{\max} \leq f_{l,t} \leq F_l^{\max}, \quad \forall l \in \mathcal{L}, t \in \mathcal{T}, \tag{6d} \\
& \Theta_n^{\min} \leq \theta_{n,t} \leq \Theta_n^{\max} \quad \forall n \in \mathcal{N}, t \in \mathcal{T}, \tag{6e} \\
& 0 \leq p_{i,t} \leq P_i^{\max}, \quad \forall i \in \mathcal{G}, t \in \mathcal{T}, \tag{6f} \\
& -RD_i \leq p_{i,t} - p_{i,t-1} \leq RU_i, \quad \forall i \in \mathcal{G}, t \in \mathcal{T}, \tag{6g} \\
& 0 \leq d_{i,t}^{nd} \leq D_{i,t}, \quad \forall i \in \mathcal{ND}, t \in \mathcal{T}, \tag{6h} \\
& 0 \leq d_{i,t}^{dc} \leq \text{gridLoad}_{i,t}, \quad \forall i \in \mathcal{DC}, t \in \mathcal{T}, \tag{6i} \\
& 0 \leq m_{i,t} \leq M_{j,t}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, \tag{6j} \\
& 0 \leq w_{i,t} \leq W_{j,t}, \quad \forall i \in \mathcal{W}, t \in \mathcal{T}, \tag{6k} \\
& 0 \leq s_{i,t} \leq S_{j,t}, \quad \forall i \in \mathcal{S}, t \in \mathcal{T}, \tag{6l} \\
& 0 \leq r_{i,t} \leq R_{j,t}, \quad \forall i \in \mathcal{R}, t \in \mathcal{T}. \tag{6m}
\end{aligned}$$

B Total Cost of Ownership (TCO) of Battery Energy Storage

Total cost of ownership (TCO) includes capital expenses (CapEx) depreciation and operational expenses (OpEx) incurred periodically.

The CapEx of battery storage can be broken down into energy component (EC) and power component (PC) costs, which correspond to batteries and electrical infrastructure, with costs proportional to the energy and power capacity respectively.

The OpEx of Li-ion batteries covers operation, maintenance, and capacity augmentation to mitigate battery degradation. We assume linear battery degradation by cycle, under which one cycle is counted when accumulated discharged energy reaches $\text{energyCap} * \text{DoD}$ (depth of discharge). Based on all above, the TCO of battery energy storage (\$/year) can be calculated by:

$$\frac{\text{power}_{\max} * \text{unitCapex}_{PC} + \text{energyCap} * \text{unitCapex}_{EC}}{\text{DeprPeriod}_{MB}}$$

$$+ \text{power}_{\max} * \text{unitCapex}_{sys} (0.01 * \text{cycle}_{avg} + 0.015) + \text{lossCost}$$

where $\text{unitCapex}_{sys} = \text{unitCapex}_{PC} + \text{unitCapex}_{EC} \cdot \text{duration}$ is a system cost metric. lossCost is calculated based on typical round-trip efficiency, daily discharge, and power price. Table 6 lists the settings used in Section 5.

Table 6: Li-ion BES Cost Settings (Sources: [13, 63])

Capex (80% DoD, 85% RTE, $\text{DeprPeriod}_{MB}=15$ years)	
Power Components	0.36 million \$/MW (2023)
Energy Components	0.39 million \$/MWh (2023)
Opex (million \$/(MW·year))	
Capacity Augmentation	$1\% \cdot \text{unitCapex}_{sys} \cdot \text{\#cycles/day}$
Others	$1.5\% \cdot \text{unitCapex}_{sys}$