Optimal Energy Procurement for Geo-distributed Data Centers in Multi-timescale Electricity Markets

Tan N. Le[1,2], Jie Liang[1], Zhenhua Liu[1],
Ramesh K. Sitaraman[3,4], Jayakrishnan Nair[5], Bong Jun Choi[1,2]
Stony Brook Univ.[1], SUNY Korea[2], Univ. of Massachusetts at Amherst[3], Akamai Tech.[4], IIT Bombay[5]

ABSTRACT

Multi-timescale electricity markets augment the traditional electricity market by enabling consumers to procure electricity in a futures market. Heavy power consumers, such as cloud providers and data center operators, can significantly benefit from multi-timescale electricity markets by purchasing some of the needed electricity ahead of time at cheaper rates. However, the energy procurement strategy for data centers in multi-timescale markets becomes a challenging problem when real world dynamics, such as spatial diversity of data centers and uncertainties of renewable energy, IT workload, and electricity price, are taken into account. In this paper, we develop energy procurement algorithms for geo-distributed data centers that utilize multi-timescale markets to minimize the electricity procurement cost. We propose two algorithms. The first algorithm provides provably optimal cost minimization while the other achieves near-optimal cost at a much lower computational cost. We empirically evaluate our energy procurement algorithms using real-world traces of renewable energy, electricity prices, and workload demand. Our empirical evaluations show that our proposed energy procurement algorithms save up to 44% of the total cost compared to traditional algorithms that do not use multi-timescale electricity markets or geographical load balancing.

1. INTRODUCTION

Data centers are becoming the largest and the fastest growing consumers of electricity in the United States. It is reported that US data centers consumed 91 billion kilowatt-hours (kWh) in 2013, which is more than twice of the electricity consumed by households in New York City (see [40]). In the same report, the electricity consumption of data centers is estimated to reach 140 billion kWh in 2020 due to the explosion of demand for cloud computing and other Internet-scale services. Global cloud providers such as Google and Amazon, who operate multiple data centers, spend billions of dollars annually on their electricity bills [33].

Multi-timescale electricity markets have been proposed to improve the efficiency of electricity markets [12]. Multi-timescale electricity markets encompass both *forward* (futures) and *spot* (real-time) markets. While energy is procured at the time of consumption in a spot market, forward markets allow customers to buy electricity a day ahead or even several months ahead of when it is consumed. Forward electricity markets reduce the risk for both the supplier and consumer by reducing the quantity of energy trading in the real-time spot markets [4]. Furthermore, purchasing electricity ahead of time can facilitate the expansion of renewable energy sources. For example, Google invested in purchasing renewable energy from renewable energy project developers for 20 years [16].

Utilizing multi-timescale markets has great potential for electricity cost savings for cloud providers who operate one or more data centers. There has been much recent work that exploits the variation of real-time electricity prices in the temporal and spatial dimensions to reduce the total electricity cost. For example, prior papers show how a cloud provider can exploit real-time electricity prices in multiple market locations and move the load to locations with a cheaper price [33, 10, 34]. Other papers exploit temporal variation in the real-time energy price and use energy storage to reduce electricity costs [18, 19, 43], i.e., the storage device is charged during the times when the electricity price is low and discharged when the price is high. However, while these works focus on traditional real-time markets, the potential of using multi-timescale markets for electricity cost reduction has not been well studied in the context of a cloud provider; this is the focus of the present paper.

In particular, using forward markets to lower the electricity cost of a cloud provider is challenging for multiple reasons. The optimal amount of electricity that a cloud provider should purchase in advance for a particular location depends on the workload, the onsite renewable generation, and the real-time electricity price at that location at the (future) time of delivery. But future workload, renewable generation, and real-time electricity price are not perfectly predictable and are subject to significant forecasting errors. Note that if the cloud provider is too conservative and buys too little from the forward market, any shortfall in electricity would need to be covered by purchasing it from the more expensive[1] real-time market. Likewise, if the cloud provider is too aggressive and buys too much from the forward market, any excess in purchased electricity will go wasted. Moreover, the ability of a cloud provider to move the load from one data center to another, possibly incurring a performance penalty that we characterize as the "delay cost", adds an additional level of complexity that needs to be optimized. In

[1]In some cases, the prices in the forward markets might be (on average) higher than real-time prices. If so, instead of saving electricity expenditure, the cloud provider can participate in forward markets to reduce cost variations. Our model can be extended to handle either case.

Performance '17 NY, USA

this work, we provide an optimization framework for tackling the aforementioned challenges.

Our contributions are three-fold.

(1) Optimal algorithm development. We develop two algorithms for a cloud provider with geo-distributed data centers to buy electricity in multi-timescale markets: one algorithm provides optimality guarantees, while the other is simpler, uses limited predictions but achieves near-optimal performance. To develop the energy procurement system, we first model the problem of procuring electricity for geo-distributed data centers in multi-timescale markets in Section 3. The system model is general and applicable to any global cloud provider with access to multi-timescale electricity markets. We focus on two-timescale markets that consist one long-term market and one real-time market, though our model and algorithms can be extended to handle multiple markets at various timescales. We present the characteristics of the objective functions and the optimal solution in Section 4, which forms the theoretical basis for our algorithm design. The two algorithms that we design, prediction based algorithm (PA) and stochastic gradient based algorithm (SGA), are described in Section 6. Both algorithms seek to minimize the total operating cost of the cloud provider across all data centers. While PA is simple and performs very well in practice, SGA provably achieves the optimal solution.

(2) Predictability analysis using real-world traces. To provide the inputs for our energy procurement algorithms, we collect and analyze real world traces of PV generation, wind generation, electricity prices, and IT workload demand. A detailed data analysis of real-world traces of PV generation, wind generation, electricity prices, and workload, is presented in Section 5. The data analysis not only enables us to evaluate our algorithms using real-world data but also provides insights into the nature of prediction errors. To procure electricity in forward markets, the energy procurement system needs to predict the renewable generation, workload, and electricity prices in real-time. Therefore, we focus on addressing the following questions. What do the distributions of prediction errors look like? How correlated are prediction errors in the spatial domain?

(3) Empirical evaluation. We carry out a detailed empirical evaluation of our proposed energy procurement systems using real world traces. In Section 7, we demonstrate that SGA can converge to the optimal solution in a small number of iterations. Moreover, we show that PA, our heuristic algorithm, surprisingly achieves a near-optimal solution. This is partially because the real-time optimization takes into consideration the trade-off between energy cost and delay cost, and is able to compensate for some prediction errors by redirection workloads. The proposed energy procurement systems are compared with other comparable energy procurement strategies to highlight their benefits. The impacts of renewable energy and prediction errors on the proposed systems are also presented.

2. BACKGROUND AND PRIOR WORK

Internet-scale services, such as Google's search services, Akamai's content delivery services, and Amazon's cloud computing services are rapidly growing, consume large amounts of energy [21]. In fact, energy costs account for a large portion of the overall operating expenditure of such services. The two main approaches for reducing the energy cost are, to procure energy in a more cost effective manner, and to reduce the total energy consumption. While there has been much work on both approaches, we provide a survey of the energy procurement literature below.

A key technique used to reduce energy costs is to exploit the *temporal* variation of energy prices and shift the delay-tolerant workload, such as batch jobs, to off-peak time periods when the electricity prices are lower [14, 25, 29, 42]. An alternate technique is to "move the energy" using a battery. By charging the batteries from the grid when the electricity prices are lower and discharging it when prices are higher, the overall energy costs can be reduced [39, 28]. Since service providers pay for the peak of electricity usage, batteries can also be used to "shave" the power peaks to reduce the energy costs [32]. The above papers all consider a single data center. In contrast, the approach in this paper is applicable to cloud providers with multiple, geographically distributed, data centers.

Another complementary technique that is relevant to service providers with multiple geo-distributed data centers is to exploit the *geographical* variation in energy prices. There has been much work in geographical load balancing (GLB) algorithms that route the workload to the regional markets with cheap electricity prices to reduce the total energy cost [33, 27]. While these works rely on the spatial diversity of electricity prices in real-time, our approach deal with the uncertainty of electricity prices in forward markets.

In addition, data centers can reduce the energy cost by utilizing onsite renewable generation. Although the output of renewable energy sources is intermittent, a single data center can schedule its delay-tolerant workloads to adapt to the renewable generation [26]. Service providers with geographically distributed data centers can even do better by shifting their workload to the data centers that have available renewable sources [9, 27, 20]. Thus, the amount of energy cost reductions heavily depends on the percentage of delay-tolerant workloads and the penetration of renewable energy.

Participating in multiple time-scale markets, i.e., in both forward electricity markets and spot markets, has not been explored in-depth in prior work, and it can be a promising approach to more effectively reduce the energy cost. The forward electricity markets, such as long-term (several months) and short-term (day-ahead), were designed to improve the traditional electricity markets, which have only spot (real-time) markets [4]. Forward electricity markets have already been adopted in some parts of the US such as New England [12]. Forward markets can benefit both customers and utility suppliers. For example, the forward markets allow suppliers and consumers to agree on a fixed price several months ahead of when the electricity is produced and consumed. This allows the supplier to plan ahead and ensure the availability of energy for its customers. The forward markets usually provide cheaper prices than the spot markets. There are a few recent papers on data centers that consider forward markets; these papers deal with the financial risk arising from the uncertainties in electricity prices and workload [35, 41]. Geographical load balancing systems with both day-head market and real-time markets has been studied in a recent publication [15]. However, the proposed solution is somewhat restrictive to particular distributions to facilitate stochastic optimization and does not provide any optimality guarantee.

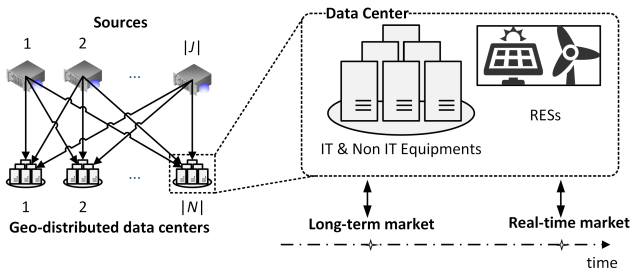


Figure 1: Geo-distributed data centers in long-term and real-time markets.

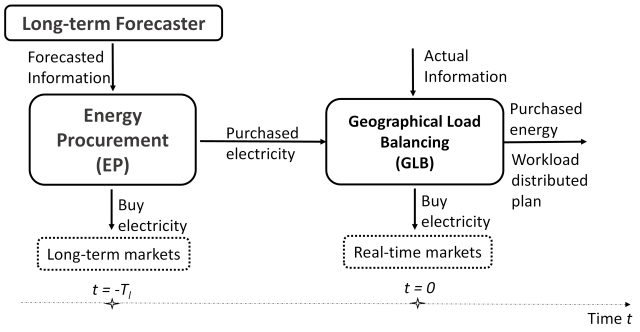


Figure 2: Energy Procurement System (EPS) Architecture for geo-distributed data centers.

3. MODEL

In this section, we present our model of the energy procurement problem for geo-distributed data centers participating in multi-timescale markets. For analytical tractability, we consider a two-timescale setting, consisting of a long-term electricity market and a real-time electricity market.

3.1 System model

Two-timescale markets. A service provider operating geo-distributed data centers can purchase electricity in two markets – a long-term market and a real-time market. The electricity consumed at time $t = 0$ must be procured from the real-time market at $t = 0$ and/or from the long term market *ahead of time* at $t = -T_l$.

Geo-distributed data centers. We consider a set N of geo-distributed data centers serving workload demands from a set J of sources as illustrated in Figure 1. The workload demand from each source is split between the $|N|$ data centers. Here, a source can represent the aggregate demand from a group of local users, such as users of a particular city, ISP, or geographical region. Each data center has access to renewable energy sources. Further, each data center participates in a (local) long-term electricity market and a (local) real-time electricity market. In other words, each data center i can buy electricity ahead of time in its long-term market, and can also buy additional electricity in its real-time market if necessary.

Energy procurement system (EPS). Our proposed energy procurement system for geo-distributed data centers is depicted in Figure 2. There are three main components, namely, the long-term forecaster, the energy procurement (EP) in long-term markets and the geographical load balancing (GLB). The long-term forecaster provides the forecasted information for the energy procurement. The forecasted information includes the predicted values and the prediction error distributions of IT workload, renewable energy generations, and electricity prices. We design the algorithms for the long-term forecaster in Section 5. The EP component procures electricity for each data center in the corresponding long-term markets (at time $t = -T_l$) based on the electricity prices in the long-term markets and *forecasts* of real-time prices, workload, and renewable generation. The GLB component (at time $t = 0$) distributes (routes) the realized workload from sources to data centers, provisions the required computing capacity at each data center, and procures additional electricity as needed in the real-time markets.

Data center. Let M_i denote the number of servers in data center i. The number of active servers at real-time (time $t = 0$) is denoted by m_i, which is a control parameter. In practice, there can be more than a hundred thousand servers in a single data center. Thus, in our mathematical modeling, we treat m_i as a real number satisfying $0 \le m_i \le M_i$.

At time $t = 0$, the power consumption of data center i is denoted by d_i^r. In general, the power consumption of data center i is dependent on the number of active servers m_i and the workload arrival λ_i. For simplicity, we assume that $d_i^r = m_i$, which implies that the power consumption is proportional to the number of active servers, and is independent of the workload λ_i.[2]

Workload. Workload demand from source j in real-time ($t = 0$) is denoted as L_j^r. We assume that the exact realization of the random vector $\boldsymbol{L}^r = (L_j^r, j \in J)$ is known to the cloud provider at time $t = 0$, and is an input to GLB. Let λ_{ij} denote the distributed workload arrival from source j to data center i at time $t = 0$ (set by GLB). Thus,

$$L_j^r = \sum_{i \in N} \lambda_{ij} \qquad (j \in J),$$

$$\lambda_i = \sum_{j \in J} \lambda_{ij} \qquad (i \in N).$$

Here, λ_i denotes the aggregate workload routed to data center i.

Renewable energy. Data centers can utilize their integrated RESs. Let w_i^r denote the renewable energy generation at data center i in real-time ($t = 0$). We assume that the exact realization of the random vector $\boldsymbol{w}^r = (w_i^r, i \in N)$ is known at time $t = 0$, and is an input to GLB.

Electricity price. For each data center, the cloud provider can purchase electricity at time $t = -T_l$ in the local long-term market and then purchase any additional electricity needed in the local real-time market at time $t = 0$. For data center i, let p_i^l denote the long-term price for 1 unit of electricity, and p_i^r denote the real-time price for 1 unit of electricity. We assume that $\boldsymbol{p}^l = (p_i^l, i \in N)$ is fixed (or equivalently, is known at the time of the long-term procurement), and $\boldsymbol{p}^r = (p_i^r, i \in N)$ is a random vector whose exact value is known is known at time $t = 0$ and is an input to GLB.

[2]The proportionality constant relating the number of active servers and the power consumption is taken to be 1 without loss of generality.

Note that the real-time workload $\boldsymbol{L}^r$, the real-time renewable generation $\boldsymbol{w}^r$, and the real-time electricity prices $\boldsymbol{p}^r$ are unknown at the time of the long-term procurement by the EP component, but are known at the time of operation of the GLB component. We assume that the random vector $(\boldsymbol{L}^r, \boldsymbol{w}^r, \boldsymbol{p}^r)$ is jointly continuous. In addition, all the w_i^r, L_j^r, and p_i^r are assumed to be bounded random variables.

3.2 Cost model

The total cost of operating geo-distributed data centers in composed of a delay cost and an energy cost. The delay cost is the monetary cost incurred due to the delay in processing the arriving workload. The energy cost is the total electricity bill from the long-term and real-time markets.

Delay cost. We consider both the network delay between data centers and sources and the processing time within a data center. The network delay π_{ij} captures the delay that propagates the workload λ_{ij} from source j to data center i. The queuing delay $g_i(m_i, \lambda_i)$ denotes the delay at data center i to process its arrival workload λ_i. For stability, we need that $\lambda_i < m_i\mu_i$. Here, μ_i is the service rate of a server in data center i. Thus, we define $g_i(m_i, \lambda_i) = \infty$ for $\lambda_i \geq m_i\mu_i$.

We model the delay cost $h_{ij}(m_i, \lambda_i)$ of routing and processing each unit of workload from source j to data center i as follows.

$$h_{ij}(m_i, \lambda_i) = \beta\big(g_i(m_i, \lambda_i) + \pi_{ij}\big). \tag{1}$$

Here, the parameter β weighs the delay relative to the energy cost. While (1) assumes a linear relationship between incurred delay and the associated monetary cost (as is suggested in [5]), our model allows for a non-linear (convex) relationship between delay and its monetary cost to the cloud provider. The delay cost of transmitting workload λ_{ij} from source j to data center i is computed as $\lambda_{ij}h_{ij}(m_i, \lambda_i)$. We assume that $h_{i,j}(m_i, \lambda_i)$ is continously differentiable over $0 \leq \lambda_i < m_i\mu_i$, and that $\lambda_{ij}h_{ij}(m_i, \lambda_i)$ is jointly convex with respect to m_i and λ_i.

A specific instance of the delay cost function h_{ij} that satisfies the above assumptions, and which we use in our experimental evaluations, is

$$h_{ij}(m_i, \lambda_i) = \beta\left(\frac{1}{\mu_i - \lambda_i/m_i} + \pi_{ij}\right) \qquad (\lambda_i < m_i\mu_i), \tag{2}$$

where the first term $\frac{1}{\mu_i - \lambda_i/m_i}$ above captures queuing delay at delay center i, which is based on the well-known mean delay formula for the M/GI/1 processor sharing queue.

Energy cost. Let q_i^l and q_i^r respectively denote the amount of electricity purchased in the long-term market and the real-time market by data center i. Here, we require that sufficient electricity is procured to process the workload routed to each data center as

$$q_i^r + w_i^r + q_i^l \geq d_i^r = m_i \quad (i \in N).$$

The electricity bills of data center i in the long-term market and the real-time market are respectively computed as

$$R_i^l(q_i^l) = p_i^l q_i^l \qquad i \in N,$$
$$R_i^r(q_i^r) = p_i^r q_i^r \qquad i \in N.$$

3.3 Formulation of optimal energy procurement in multi-timescale markets

In this section, we describe the optimization formulation for optimal energy procurement. Recall that the total cost of operating geo-distributed data centers in our two-timescale market setting is the sum of the energy cost and the delay cost, given by

$$F = \sum_{i\in N} R_i^l(q_i^l) + \sum_{i\in N} R_i^r(q_i^r) + \sum_{i\in N, j\in J} \lambda_{ij}h_{ij}(m_i, \lambda_i).$$

We seek to minimize $\mathbb{E}[F]$ subject to the aforementioned constraints. Note that this optimization is performed on two timescales, with different sets of information available at each. The EP component optimizes the long-term procurements $\boldsymbol{q}^l = (q_i^l, i \in N)$ given only distributional information of the real-time workload $\boldsymbol{L}^r$, the real-time renewable generation $\boldsymbol{w}^r$, and the real-time electricity prices $\boldsymbol{p}^r$. The GLB component optimizes the workload routing $\boldsymbol{\lambda} = (\lambda_{ij}, i \in N, j \in J)$, the number of active servers $\boldsymbol{m} = (m_i, i \in N)$ at the data centers, and the real-time procurements $\boldsymbol{q}^r = (q_i^r, i \in N)$ given the prior long-term procurements $\boldsymbol{q}^l$, and the exact realization of $(\boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$. Below, we first formalize the real-time optimization, followed by the long-term optimization.

Geographical load balancing in real-time markets. Note that in real-time, GLB optimizes the real-time procurements $\boldsymbol{q}^r$, the numbers of active servers $\boldsymbol{m}$, and the workload routing $\boldsymbol{\lambda}$, given the long-term procurements $\boldsymbol{q}^l$ and the realization of the random vector $(\boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$. The total cost as seen by GLB is

$$F^r(\boldsymbol{q}^r, \boldsymbol{m}, \boldsymbol{\lambda}, \boldsymbol{p}^r) := \sum_{i\in N} R_i^r(q_i^r) + \sum_{i\in N, j\in J} \lambda_{ij}h_{ij}(m_i, \lambda_i).$$

Thus, the real-time optimization is defined as follows.

$$\textbf{GLB-RT:} \min_{\boldsymbol{m},\boldsymbol{\lambda},\boldsymbol{q}^r} F^r(\boldsymbol{q}^r, \boldsymbol{m}, \boldsymbol{\lambda}, \boldsymbol{p}^r)$$

s.t.

$$\lambda_{ij} \geq 0 \quad \forall i \in N, j \in J \tag{3a}$$
$$\sum_{i\in N} \lambda_{ij} = L_j^r \quad \forall j \in J \tag{3b}$$
$$\lambda_i \leq m_i\mu_i, \quad \forall i \in N \tag{3c}$$
$$0 \leq m_i \leq M_i \quad \forall i \in N \tag{3d}$$
$$q_i^r \geq 0, \quad \forall i \in N \tag{3e}$$
$$m_i - q_i^r - w_i^r \leq q_i^l \quad \forall i \in N \tag{3f}$$

Since $p_i^r \geq 0$, it easily follows that any solution of the above optimization problem satisfies $q_i^r = [m_i - w_i^r - q_i^l]_+$, where $[x]_+ := \min\{0, x\}$. Thus, the real-time objective can be re-written as

$$\begin{aligned}\tilde{F}^r(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda}, \boldsymbol{p}^r, \boldsymbol{w}^r) = {} & \textstyle\sum_{i\in N} p_i^r[m_i - w_i^r - q_i^l]_+ \\ & + \textstyle\sum_{i\in N, j\in J} \lambda_{ij}h_{ij}(m_i, \lambda_i).\end{aligned} \tag{4}$$

With this notation, GLB-RT can be equivalently expressed as follows.

$$\min_{\boldsymbol{m},\boldsymbol{\lambda}} \tilde{F}^r(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda}, \boldsymbol{p}^r, \boldsymbol{w}^r)$$

s.t.

$$(\boldsymbol{m}, \boldsymbol{\lambda}) \in C(L^r).$$

Here, the convex compact set $C(L^r)$ is defined by the constraints (3a)–(3e).

GLB-RT problem is a convex optimization problem and hence can be solved efficiently using standard techniques [7]. For instance, CVX (Matlab Software for Disciplined Convex Programming) tool [17] can be used to solve GLB-RT.

Energy procurement in long-term markets. At time $t = -T_l$, the cloud provider purchases electricity $\boldsymbol{q}^l$ in long-term markets that will be used at real-time. Note that optimization of the long-term procurements has to be performed based only on distributional information for the random vector $(\boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$, and subject to the real-time optimization that will be subsequently performed based on the realization of the random vector $(\boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$.

Let us denote the optimal value of the optimization GLB-RT by $F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$. The long-term objective is thus defined as

$$F^l(\boldsymbol{q}^l) := \sum_{i \in N} R_i^l(q_i^l) + \mathbb{E}\left[F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)\right].$$

Note that the above expectation is with respect to the random vector $(\boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$. The long-term optimization problem is then given by:

$$\textbf{EP-LT:} \ \min F^l(\boldsymbol{q}^l)$$
$$\text{subject to}$$
$$\boldsymbol{q}^l \in \mathbb{R}_+^N.$$

The above optimization is more challenging than GLB-RT. In Section 4, we prove that EP-LT is a convex optimization and characterize the gradient of the objective function. These results are then used to arrive at a provably optimal stochastic gradient algorithm in Section 6.

4. CHARACTERIZING THE OPTIMA

In this section, we collect useful properties of the optimizations EP-LT and GLB-RT. These are important for understanding the behavior of the energy procurement system, and also for proving convergence of the stochastic gradient algorithm for EP-LT in Section 6.

Our first result is that EP-LT is indeed a convex optimization, which suggests that EP-LT is a tractable optimization.

THEOREM 1. *$F^l(\boldsymbol{q}^l)$ is convex over $\boldsymbol{q}^l \in \mathbb{R}_+^N$.*

We provide the proof of Theorem 1 in Appendix A.1. Next, we characterize the gradient of the EP-LT objective function as follows.

THEOREM 2. *The gradient of $F^l(\cdot)$ is characterised as follows.*

$$\frac{\partial F^l(\boldsymbol{q}^l)}{\partial q_i^l} = p_i^l + \mathbb{E}\left[\frac{\partial F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)}{\partial q_i^l}\right]$$
$$= p_i^l - \mathbb{E}\left[\varrho_i(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)\right],$$

where $\varrho_i(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$ is the unique Lagrange multiplier of GLB-RT corresponding to the constraint (3f).

Note that the first equality in the theorem statement asserts that the order of an expectation and a partial derivative can be interchanged. The second equality relates the partial derivative of F^{*r} with respect to q_i^l to a certain Lagrange multiplier of GLB-RT. We provide the proof of Theorem 2 in Appendix A.2.

We note that Theorem 2 does not enable us to compute the gradient of the $F^l(\cdot)$ exactly. Indeed, the expectation of the Lagrange multiplier ϱ_i with respect to $(\boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$ would in general be analytically intractable. However, Theorem 2 does enable a noisy estimation of the gradient of the $F^l(\cdot)$ via Monte Carlo simulation as follows. Suppose we simulate a finite number, say $\mathbb{S}$, of samples from the distribution of $(\boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$. In practice, we can obtain these samples by using real-world traces as is done in Section 7. For each sample, the Lagrange multipliers $(\varrho_i, i \in N)$ can be computed efficiently by solving GLB-RT. By averaging the $\mathbb{S}$ instances of $(\varrho_i, i \in N)$ thus obtained, we get an unbiased estimate of the gradient of $F^l(\cdot)$. This, in turn, enables us to solve EP-LT using a stochastic gradient descent method; details follow in Section 6.

As there are two timescales in optimization, it is critical to investigate how EP-LT affects the operation of geographical load balancing in real-time. We start by answering the following question: how does the long-term procurement $\boldsymbol{q}^l$ impact the power consumption d_i^r in data center i? Formally, we have the following intuitive result:

LEMMA 3. *At any data center i, an optimal solution always utilizes the long term energy procurement q_i^l and renewable generation w_i^r as much as possible. It is simply represented by*

$$\begin{cases} m_i \geq w_i^r + q_i^l & \text{if } w_i^r + q_i^l < M_i \\ m_i = M_i & \text{if } w_i^r + q_i^l \geq M_i. \end{cases} \tag{7}$$

PROOF. Appendix A.3. □

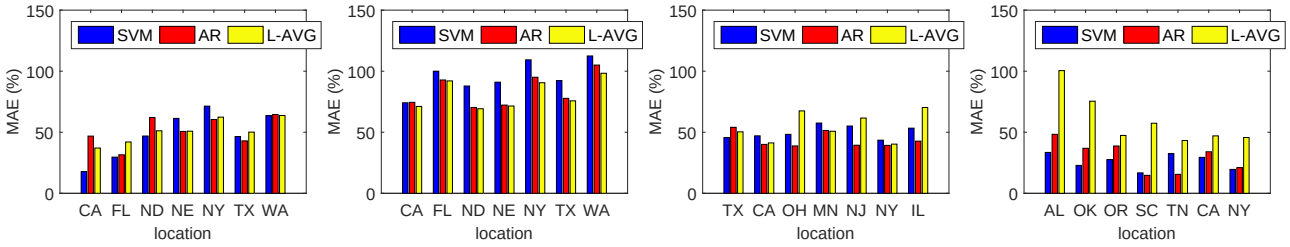
The above lemma states that a data center i uses up the reserved electricity, including free renewable energy and prepurchased electricity, because doing so reduces the queueing delay.

5. PREDICTABILITY ANALYSIS

In this section, we study the long-term predictability of metrics critical to our procurement systems for multi-timescale markets; namely, workload, renewable generation, and real-time electricity price. We design two long-term prediction methods, and analyse the prediction errors associated with each metric using real-world traces. Our analysis also provides several insights into the nature of the *distributions* of these metrics.

For our study, we collected 3-year real-world traces of photovoltaic (PV) generation, wind generation, and electricity prices for 20 states of the US. The 3-year PV and wind generation data were downloaded using the System Advisor Model (SAM) software, developed by the National Renewable Energy Laboratory (NREL) [1]. The 3-year electricity price data are from different regional transmission operators (RTOs) in the US, i.e., PJM, MISO, CAISO, ISONE, and NYSIO [33]. In addition, we collected 2-month workload data for the same 20 states from Akamai Technologies, which serves 15-30% of all Web content around the world from hundreds of data centers around the world [31].

Long-term prediction is challenging for both statistical and physical prediction methods [23]. Statistical methods have to deal with the weak correlation between the past and future data. Meanwhile, physical methods require the input



(a) PV generation (b) Wind generation (c) Electricity price (d) Workload

Figure 3: Comparisons of SVM, AR and L-AVG. The codes of US states are California (CA), Florida (FL), North Dakota (ND), Nebraska (NE), New York (NY), Texas (TX), Washington (WA), Ohio (OH), Minnesota (MN), New Jersey (NJ), Illinois (IL), Alabama (AL), Georgia (GA), Oklahoma (OK), South Carolina (SC), Virginia (VA), and Tennessee (TN).

of physical features that are often not available for long-term predictions. For example, long-term weather forecast requires data from many parts of the world which are only available in some specialized centers. To improve the prediction accuracy, prediction methods may exploit seasonality, such as annual patterns. However, the effectiveness of using seasonality depends heavily on the characteristics of the data.

We design two long-term prediction methods to produce the inputs for our energy procurement system: An autoregressive (AR) model and a Support Vector Machine (SVM) model. The motivation for using the AR method is to capture daily patterns and the correlation between past and future data. On the other hand, we develop the SVM method to capture the seasonality of the data.

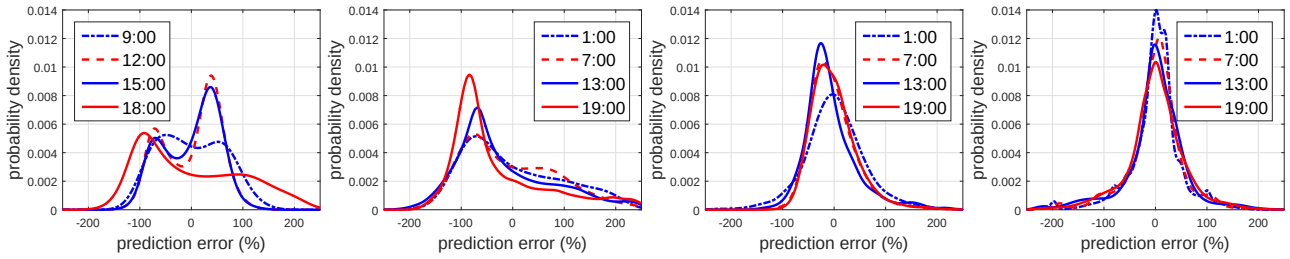
In particular, our AR model predicts the value $x(day + d\_ah, hr)$ at hour hr for d_ah day-ahead based on the past A days as $x(day + d\_ah, hr) = \sum_{a=0}^{A-1} \omega_a x(day - a, hr) + c$. The AR model can obtain the coefficients ω_a and constant c by fitting the model to the historical data. We observe that it is not necessary to pick a large value of A for long-term prediction because $A = 7$ already achieves competitive performance. Additionally, d_ah is set at 30 days for PV generation, wind generation, and electricity price, and at 1 day for workload due to the limited length of data.

Our SVM model is designed to capture the seasonality of workload, renewable generation, and electricity price. Similar to the work [37], we use a multi-class SVM. The first input to the SVM model is the average of the past A days. The rest of inputs are the seasonality data, i.e., month of year, day of month, day of week, and hour of day [13]. For electricity generation from PV panels and wind turbines, we use month of year, day of month, and hour of day to capture their seasonality. Similarly, we use month of year, day of month, hour of day, and day of week in predicting electricity prices. Due the limitation of the trace length, only day of week and hour of day are used as the seasonality inputs for predicting workload. The prediction window is the same as with the AR method, i.e., 30 days for solar generation, wind generation, and electricity prices, and 1 day for workload. The accuracy of SVM depends on the selection of SVM kernel function and the kernel parameters. For each set of data, we search for the best kernel function and the best kernel parameters using LIBSVM, an SVM tool [8]. The most suitable kernel function is Radial Basis Function (RBF) but the kernel parameters differ for each dataset.

Prediction error analysis: We now analyse the prediction errors under the AR and SVM methods. We normalize the prediction errors by the average of real values and show the values in percentage. For instance, a prediction error for PV generation of 20 (-20) implies that we underestimate (overestimate) the PV generation by 20% of the average PV generation.

We compare AR and SVM with a baseline method, long-term average (L-AVG) [38]. L-AVG assumes that the long-term data has a long-term cycle. For example, PV generation may have a yearly cycle. L-AVG takes the average of 30 days at the same time over the past 2 years for PV generation, wind generation, and electricity price. In particular, the predicted value at hour hr of day day in yr is computed as $x(yr, day, hr) = \frac{1}{2}\sum_{y=1}^{2}\frac{1}{30}\sum_{d=15}^{-14} x(yr - y, day - d, hr)$. Assuming that user behavior has a weekly pattern, L-AVG takes the averages of the workload demand at the same time of 7 days in the past weeks. The mean absolute errors (MAEs) of three methods are illustrated in Figure 3. In general, SVM and AR do not perform better than L-AVG in predicting solar and wind generation but they are better than L-AVG in predicting electricity price, and workload. In predicting PV generation, SVM outperforms others methods in some states like California that reveal the positive impact of seasonality. However, some states like Washington (WA) and New York (NY) having high precipitation can negatively affect the performance of SVM and AR. Predicting wind generation is the hardest among the four types of data as the prediction errors are very large. It is because the wind generation often has very large variation and fluctuation. On the other hand, SVM and AR are better than L-AVG in terms of predicting electricity prices. AR is surprisingly better than SVM in most states except for Texas (TX). The seasonality in real-time electricity prices is not strong enough to benefit SVM in long-term prediction. AR and SVM perform very well in predicting workload compared to L-AVG in short-term. Overall, the long-term prediction errors are relatively large compared to the mean of measured data. Figure 3 also highlights that long prediction errors are dependent on locations, the types of data, and the prediction methods.

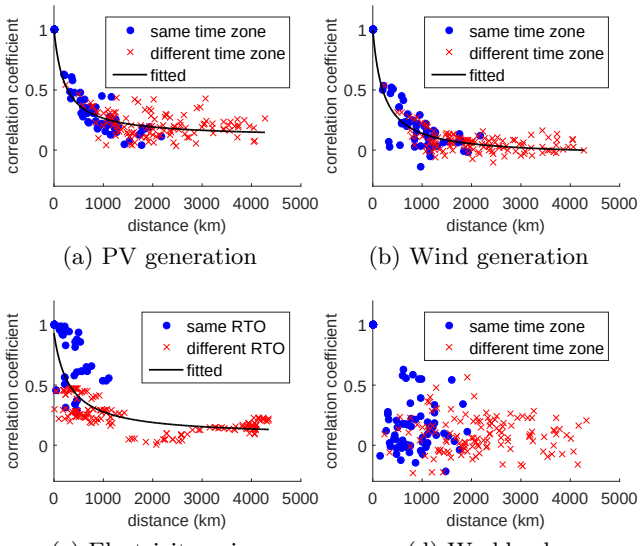
What do the distributions of prediction errors look like? Figure 4 shows the probability density of the prediction errors at different times in a day of using the AR method. Each line represents the probability density of prediction errors during the same hour for all days. The prob-



(a) PV generation (b) Wind generation (c) Electricity price (d) Workload

Figure 4: Probability density of prediction errors at different time of the day.

ability densities are obtained by averaging the probability densities of all the collected data. Our first observation is that prediction errors have zero-mean. However, the probability densities of PV generation, wind generation, electricity price, and workload are asymmetric. In particular, our prediction algorithms tend to over-predict wind generation with high probability as shown by the peaks around -80 in Figure 4(b). This is because wind generation is often low. Meanwhile, the peaks of electricity price prediction errors are close to zero-mean. The prediction errors of workload are more around zero.



(a) PV generation (b) Wind generation

(c) Electricity prices (d) Workload

Figure 5: Correlation coefficients of prediction errors in spatial domain.

How correlated are the prediction errors in spatial domain? The correlation of prediction errors in the spatial domain is of great interest to cloud providers with geo-distributed data centers. The correlation coefficients of prediction errors using AR with respect to the distance between two locations are shown in Figure 5. We classify PV generation, wind generation, and workload into two groups: within the same time zone or different time zones. There are also two groups of electricity prices: within the same RTO or different RTOs. Figure 5 highlights that both distances and groups can have the great impact on the correlation of prediction errors. The prediction errors of PV and wind generation are strongly correlated (greater than 0.5) within 500 km, weakly correlated (less than 0.5) within 1000 km, and almost independent of each other when more than 1500 km apart. Note that electricity price is more correlated in the spatial domain than PV generation and wind generation due to the fact that some of the prices can be generated by the same RTO. However, the prediction errors of workload are uncorrelated with respect to distances and groups. This is because the workload depends on unpredictable user behavior and the dynamic Internet conditions.

6. ALGORITHM DESIGN

The energy procurement system needs algorithms for both energy procurement in long-term (EP-LT) and geographical load balancing in real-time (GLB-RT). GLB-RT is a convex optimization problem that can be solved efficiently in real-time by standard techniques [27]. Thus, we focus on designing algorithms for energy procurement in the long-term markets. Note that even though EP-LT is a convex optimization (see Theorem 1), neither the objective function nor its gradient admit a closed-form representation, which presents significant challenges.

In this section, we design two algorithms, namely, Prediction based Algorithm (PA) and Stochastic Gradient estimate based Algorithm (SGA) for solving EP-LT. PA is a heuristic algorithm that requires only the predicted values of renewable generations, workload, and electricity prices. On the other hand, SGA comes with a convergence guarantee, but requires *samples* from the joint distribution of renewable generations, workload, and electricity prices. As a result, SGA can be solved in a *data-driven* manner.

6.1 Prediction based Algorithm (PA)

Prediction based algorithm (PA) relies on the mean values of renewable generation, workload, and electricity price. Fortunately, our data analysis reveals that our prediction errors for these quantities are approximately zero mean. Thus, the predicted values $\hat{L}_j^r$, $\hat{w}_i^r$, and $\hat{p}_i^r$ are good estimates of the mean values of renewable generation, workload, and electricity price.

PA computes the long-term procurement $\boldsymbol{q}^l$ by solving EP-LT and GLB-RT at the same time, with the random variables w_i^r, L_j^r, and p_i^r replaced by their predicted values. Formally, this is done by solving the following deterministic

convex optimization problem.

$$\text{LT-PA: } \min_{\mathbf{m},\boldsymbol{\lambda},\mathbf{q}^l} \sum_{i=1}^{N} p_i^l q_i^l + \sum_{i=1}^{N} \hat{p}_i^r [m_i - \hat{w}_i^r - q_i^l]_+ + \beta \sum_i \sum_j h_{ij}(m_i, \lambda_{ij})$$

subject to

Constraints (3a), (3c)–(3e)

$$\sum_{i \in N} \lambda_{ij} = \hat{L}_j^r \quad \forall j \in J$$

$$q_i^l \geq 0 \ \forall i \in N$$

The objective function of LT-PA is similar to that of the EP-LT without the expectation operation. The constraints over $\boldsymbol{m}$, $\boldsymbol{\lambda}$, and $\boldsymbol{q}^l$ of LT-PA are identical to those of GLB-RT and EP-LT. LT-PA is a convex optimization problem and can be solved efficiently by standard techniques [7]. Even though PA is a heuristic, our experimental evaluations reveal that it provides a near-optimal solution in realistic scenarios; see Section 7.

6.2 Stochastic Gradient-based Algorithm (SGA)

Although PA can offer a quick heurictic decision, it is desirable to have an algorithm that optimally procures electricity in long-term markets. To this end, we exploit the gradient characterization of the long-term objective (see Theorem 2) to design a stochastic gradient descent algorithm. The algorithm, namely, SGA, is summarized in Algorithm 1. The main idea of the algorithm is to compute a noisy estimate of the gradient of the long-term objective by averaging the gradient of the (random) total cost over a finite number of sample paths. This noisy gradient is used to perform a stochastic gradient descent. Stochastic approximation theory can then be used to prove convergence to the set of optimal solutions, as long as the step-size sequence is appropriately diminishing [22].

Algorithm 1 Stochastic Gradient based Algorithm (SGA).

Input: Obtain $\boldsymbol{p}^l$ from the $|N|$ long-term electricity markets.
Prepare $\mathbb{S}$ samples of $(\boldsymbol{w}^r, \boldsymbol{L}^r, \boldsymbol{p}^r)$ based on prediction error distributions.
Output: $q_i^l \ \forall i \in N$
Initialize: $q_i^l = 0$, $\forall i \in N$.
Step: $\tau = 1$.
while true **do**
  for all k such that $1 \leq k \leq \mathbb{S}$ **do**
    Solve: GLB-RT for kth sample of $(\boldsymbol{w}^r, \boldsymbol{L}^r, \boldsymbol{p}^r)$ with long-term procurement $\boldsymbol{q}^l$
    Obtain: The Lagrange multipliers $\varrho_i^{(k)}$ corresponding to constraint (3f), $\forall i \in N$
  end for
  Compute: $\hat{\varrho}_i = \frac{1}{\mathbb{S}} \sum_{k=1}^{\mathbb{S}} \varrho_i^{(k)}$, $\forall i \in N$
  Update: $q_i^l = [q_i^l - \eta_\tau (p_i^l - \hat{\varrho}_i)]_{[0,M_i]}$ for $\forall i \in N$. $[z]_{[0,M_i]}$ indicates the projection of z onto the set $[0, M_i]$.
  Increase: $\tau = \tau + 1$.
end while

We prove that SGA converges to the set of optimal solutions of EP-LT under the following standard assumption on the step-size sequence.

ASSUMPTION 1. $\sum_{\tau=1}^{\infty}(\eta_\tau) = \infty$ *and* $\sum_{\tau=1}^{\infty}(\eta_\tau)^2 < \infty$.

The convergence of SGA is asserted by the following theorem.

THEOREM 4. *Under Assumption 1, almost surely, the iterates* $\boldsymbol{q}^l$ *generated by SGA converge to the set of optimal solutions of EP-LT as* $\tau \to \infty$.

We give the proof of Theorem 4 in Appendix B

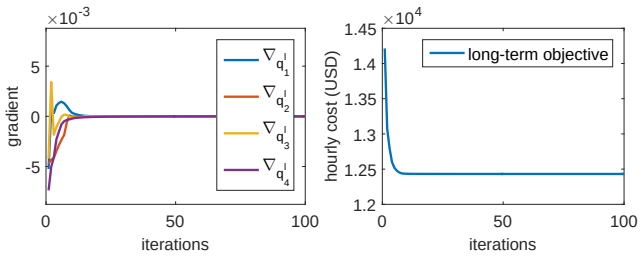
Note that SGA requires samples from the joint distribution of $(\boldsymbol{w}^r, \boldsymbol{L}^r, \boldsymbol{p}^r)$. This means that SGA can be solved in an entirely data-driven manner, without needing to actually model the distributions of workload, renewable generation, and electricity price, or the complex inter-dependencies between these quantities. This makes it particularly suitable in today's 'big-data' era. The bottleneck of SGA is the computation of the noisy gradient estimate, which involves solving $\mathbb{S}$ instances of GLB-RT. Moreover, the diminishing step-size sequence implies that SGA requires a large number of iterations to compute a near-optimal solution. However, it is important to note that since this algorithm is only used for long-term procurement, its computation time would not be a bottleneck in practice.

7. EMPIRICAL EVALUATION

Experimental Setup. There are 14 data centers in our system. They are located in 10 different states known to have Google data centers: California, Washington, Oregon, Illinois, Georgia, Virginia, Texas, Florida, North Carolina, and South Carolina. We merged the data centers in each state creating 10 logical data centers in our simulation, i.e., $|N| = 10$. We assume that there are one million servers distributed across the ten logical data centers, which is around half of the number of servers in Amazon Web Services (AWS) [30]. We take the power consumption of each server to be 300W, and consider the M/G/1 delay model (2). We consider 40 sources, corresponding to 40 states of the US; the corresponding workload data is obtained from Akamai Technologies. We use the model (2) for capturing the monetary cost of delay. The average workload is 30% of the total capacity of the data centers. The network delays π_{ij} are estimated to be proportional to the distance between sources and data centers [3]. The average network delay is 22 ms. The parameter β is estimated according to the fact that 100 ms latency costs 1% of Amazon in sales [24].

To compute the energy costs of the system, we assume that the system purchases energy in long-term markets and real-time markets for an hour of operation, i.e., the time horizon is one hour. The electricity prices in real-time markets are the industrial electricity prices of each state in May 2010 [2]. Specifically, the mean values of real-time electricity prices, $\mathbb{E}[p_i^r]$, of the considered states (in cents per kWh) are as follows: 10.41 in California, 3.73 in Washington, 5.87 in Oregon, 7.48 in Illinois, 5.86 in Georgia, 6.67 in Virginia 6.44 in Texas, 8.60 in Florida, 6.03 in North Carolina, and 5.49 in South Carolina. Since electricity prices in long-term markets are usually much cheaper than that of the real-time markets, we set the long-term prices such that the ratio $\frac{\mathbb{E}[p_i^r]}{p_i^l} = 2.5$ for all results, except for Figure 9, where the ratio is varied.

To simulate the uncertainties, the error distributions at 13 pm (like Figure 4) are used to generate the samples of



(a) Gradient components (b) Long-term objective

Figure 6: Convergence analysis.

renewable energy generation (PV generation and/or wind generation), workload, and electricity price. The mean absolute errors (MAE) of prediction errors for PV generation, wind generation, electricity price, and workload demand are 45%, 65%, 40%, and 35%, respectively. The MAE are varied later to study the impacts of prediction errors. Wind generation is used as the renewable energy source by default. The penetration of the renewable energy is fixed at 50% of the averaged demand. We also vary the penetration of PV and wind generation to investigate the impacts of the renewable portfolio and penetration level.

Convergence of SGA. Although SGA is proved to eventually converge to the optimal value of EP-LT, the convergence can be slow in practice. The convergence speed mainly depends on how the step sizes are set. Stochastic approximation is known to have high computational complexity due to the large numbers of iterations and samples needed for each iteration. To reduce the number of iterations, we use the step size update rule as $\eta_t = \frac{s}{(S+t+1)^\alpha}$, where s and S are non-negative constants and $0.5 < \alpha \leq 1$. This form fulfills the requirement of Assumption 1. To speed up the convergence of algorithm, each gradient component has its own step-size, and the step-size is updated only if the gradient component switches from negative to positive or vice versa. Figure 6 illustrates four gradient components (of total ten) and the long-term objective function updated over iterations. As shown in the figure, gradient components, and the long-term objective $F^l(q^l)$ converge very quickly, i.e it is very close to the optimal value after merely 20 iterations. In general, some gradient components $\nabla_{q_i^l}$ may converge to positive values. In such cases, the optimal solution has $q_i^l = 0$.

Cost savings. We highlight the benefit of our proposed system by comparing with the following algorithms.

No long-term procurement or geographical load balancing (nLTnGLB): nLTnGLB does not participate in long-term markets, i.e. $q_i^l = 0$, $\forall i \in N$, and the workload demand are forwarded to the closest data centers, a.k.a., the nearest routing method. We assume that the data centers activate all servers to minimize the queueing delay, i.e. $m_i = M_i$. Though simple, this policy is still widely used in practice.

Fixed long-term procurement without geographical load balancing (fLTnGLB): Cloud providers purchase a fixed amount of electricity ahead. We assume that the long-term procurement is 50% of workload mean. Like nLTnGLB, it uses the nearest routing method instead of GLB-RT.

No long-term procurement but with geographical load balancing (nLT): In this algorithm, cloud providers do not purchase the energy in long-term markets like nLTnGLB. However, they execute GLB-RT to minimize the total cost in real-time.

Fixed long-term procurement geographical load balancing (fLT): fLT buys a fixed amount of electricity in long-term markets same as fLTnGLB, i.e., 50% of workload mean. In real-time markets, it executes GLB-RT.

In addition to the baseline algorithms, we compare our algorithms to *Oracle Algorithm (OA)*. OA is an unrealizable algorithm that is given to the absolute performance limit by assuming all realizations of renewable energy, workload, and electricity prices are fully known apriori. Similarly to PA, the problem of long-term procurement can then be solved efficiently. The cost of OA is measured by averaging its output over many realizations.

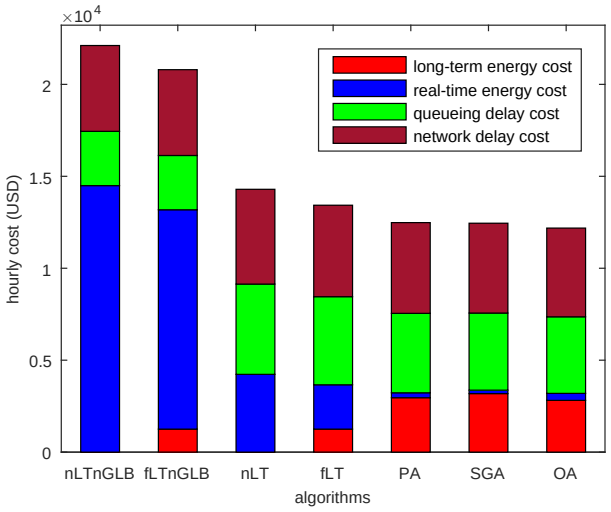
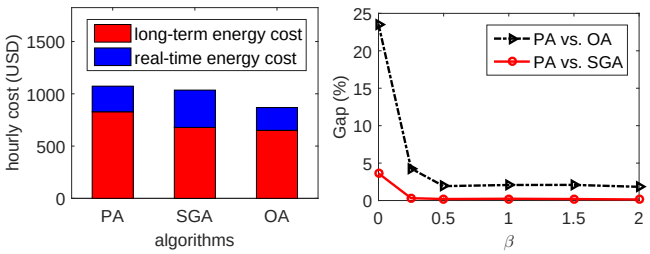


Figure 7: Cost comparison when $\beta = 1$ and 50 % renewable penetration. The proposed algorithms PA and SGA are very close to the lower bound, OA and outperform the traditional methods up to 44%.

Figure 7 compares the one-hour costs among our proposed algorithms and the traditional algorithms. The figure highlights that our proposed algorithms PA and SGA save up to 44% compared to other simpler algorithms, and are comparable to the oracle algorithm (OA), the impractical lower bound. It also shows the significant benefits for cloud providers to participate in long-term markets. Surprisingly, the performance of PA is very close to that of the SGA.

Why do our proposed algorithms perform so well? The intuition behind the small performance gaps between PA, SGA and OA is the compensation of GLB-RT at real-time markets. In particular, GLB-RT can utilize the available renewable energy and cheap electricity to partially compensate for performance gap caused by the prediction errors in long-term. More interestingly, PA and SGA are noticeably aggressive in long-term markets as in Figure 7. In addition, PA and SGA are even more aggressive than OA. Note that additional energy procured in the long-term market can be utilized to reduce queueing delay in real-time. Thus, there is the trade-off between the energy costs and delay costs that helps our proposed methods become close to OA.

How does the trade-off between energy costs and delay costs benefit our proposed algorithms? To answer this ques-

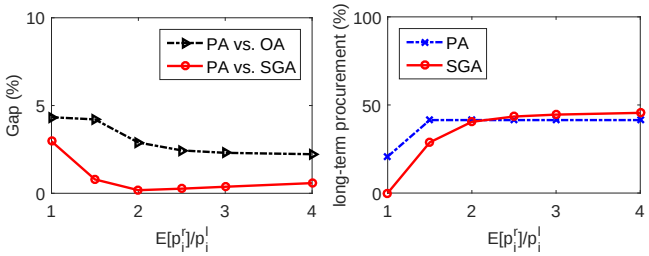


(a) $\beta = 0$ (b) Vary β

Figure 8: The impact of delay on the proposed algorithms.

tion, we vary the constant factor β that weighs the delay costs relative to energy costs. When $\beta = 0$, i.e., the delay costs are ignored, the cost breakdown are shown in Figure 8a. The performance gap between PA and OA is 24% that is much larger than the 2% gap in Figure 7 ($\beta = 1$). In this setting, SGA outperforms PA by 4%. We observe that PA is more aggressive compared to SGA in long-term procurement. Figure 8b shows the performance gaps of PA versus OA and PA versus SGA with varying β. In this figure, the x-axis shows a scaled β, where a value of 1 corresponds to the default value. We note that the performance gaps are significant when β is small (< 0.25). However, the gaps are very small when β is relatively large (≥ 0.5).

Sensitivity Analysis. The capability of our proposed algorithms depends on multiple impact factors, such as the ratio of real-time price to long-term price, renewable penetration rates, and prediction errors.

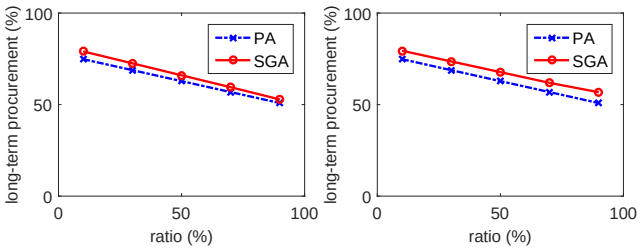


(a) PA vs. OA (b) Long term procurement

Figure 9: The impacts of long-term prices on the proposed algorithms. The gaps between the proposed algorithms and OA are small at various ratios of real-time prices to long-term prices.

Impact of the ratio of real-time price to long-term price. We carry out another study that quantifies the impact of the ratio of real-time prices to the long-term prices on our proposed algorithms as in Figure 9. In this experiment, the long term prices are fixed, and we scale the real-time prices. Figure 9a shows the performance gaps of PA versus SGA and PA versus OA. In general, the gaps are small whatever the ratio is. Figure 9b illustrates the behaviors of PA and SGA in long-term markets. SGA is more conservative than PA when the ratio is small (< 2). When the real-time prices are as cheap as the long-term prices, being more aggressive in long-term actually results in higher financial risk to the cloud providers. In contrast, SGA is more aggressive in long-term markets as the ratio becomes larger than 2.

Impact of renewable energy. Renewable energy has been increasingly used to power data centers. Hence, we investigate the impacts of renewable energy integration on our energy procurement system. We scale the penetration levels of renewable energy from 5% to 95% of the total demand. We consider PV generation and wind generation as two main sources of renewable energy.

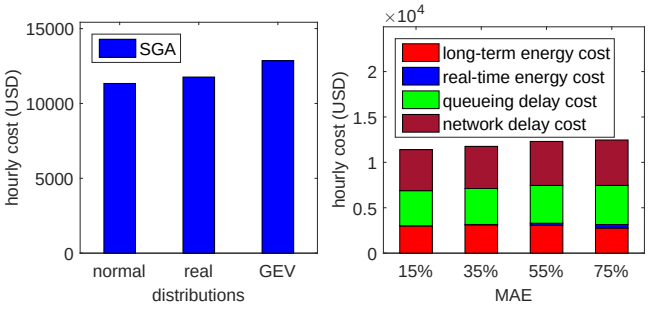


(a) PV penetration (b) Wind penetration

Figure 10: Impacts of renewable energy penetration levels on long-term energy procurement. SGA becomes less aggressive in the PV generation case than the wind generation case compared to PA.

The impacts of renewable energy on the behaviors of PA and SGA are shown in Figure 10. Here, the x-axis represents the penetration levels of renewable energy, and the y-axis is the ratio (%) of total electricity purchased in long-term markets. PA performs similarly in both cases because it is only based on the predicted values. However, SGA is closer to PA in the PV generation case as the penetration of renewable energy increases, yet becomes more aggressive than PA in the wind generation case. The reason lies in the error distributions in Figure 4. While the prediction errors of PV generation are concentrated on two peaks, the prediction errors of wind generation are centered around only one peak (around -80%).

Impact of prediction errors. So far, we have worked with the empirical (or 'real') prediction error distributions. We now study the dependence of the distribution of prediction errors on the performance of our procurement system.



(a) Error distributions (b) Prediction errors

Figure 11: Impacts of predictions on cost performance.

Figure 11a presents the cost of SGA under three different error distributions, i.e. normal, 'real', and generalized extreme value (GEV) type I (Gumbel) [11]. The normal distribution is symmetric around its mean. The GEV distribution is asymmetric and widely used in risk management and finance. We also consider the distribution of AR prediction errors as the 'real' distribution. The MAEs of each

are set at 35% for fair comparison. Figure 11a shows that the cost using normal distribution is the best among three error distributions while GEV is the worst.

Figure 11b shows the cost of SGA with respect to different MAEs of 'real' distribution. As the prediction errors increase, the real-time cost (real-time energy cost and delay cost) increase to compensate for the mis-provisioning in long-term markets. Furthermore, the total cost increases by 10% as the prediction errors increase from 15% to 75%.

8. CONCLUDING REMARKS

In this paper, we present a systematic study of optimal energy procurement systems for geo-distributed data centers that utilize multi-timescale electricity markets. The contributions of this paper are three-fold: (i) designing algorithms for long-term electricity procurement in multi-timescale markets; (ii) analyzing long-term prediction errors using real-world traces; and (iii) empirically evaluating the benefits of our proposed procurement systems. In particular, we proposed two algorithms, PA and SGA, both of which save up to 44% of the energy procurement cost compared to traditional algorithms that do not use long-term markets or geographical load balancing. While SGA provably converges to an optimal solution, PA surprisingly achieves a cost that is nearly optimal with much less computing effort.

There are a number of interesting directions for future research that can be motivated by our work. For example, generalizing our model to include more complicated forward contracts that procure energy that can be used over multiple time-slots is also another challenging problem. Integrating storage capabilities, e.g., batteries and/or thermal storage, into the energy procurement optimization of multi-timescale markets is another challenging direction. The further research on how to optimally utilize multi-timescale markets will have high potential to greatly impact the cost efficiency of Internet-scale services.

9. ACKNOWLEDGEMENTS

This research is supported by NSF grants CNS-1464388, CNS-1617698, CNS-1730128, CNS-1717588, and CNS-1413998. This research was partially funded by the MSIP, Korea, under the "ICT Consilience Creative Program" (IITP-2015-R0346-15-1007) supervised by the IITP and under the "Basic Science Research Program" (NRF-2015R1C1A1A01053788) through the NRF, Korea.

10. REFERENCES

[1] `https://sam.nrel.gov/`, 2010.

[2] `http://www.eia.gov/`, 2015.

[3] AT&T. U.S. network latency, 2016.

[4] L. M. Ausubel and P. Cramton. Using forward markets to improve electricity market design. *Utilities Policy*, 18(4):195–200, 2010.

[5] H. Beheshti and A. Croll. Performance impact: How web speed affects online business KPIs. In *Velocity Online Conference*. O'Reilly, 2009.

[6] D. P. Bertsekas, A. Nedic, and A. Ozdaglar. *Convex analysis and optimization*. Athena Scientific, 2003.

[7] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[8] C.-C. Chang and C.-J. Lin. LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3):27, 2011.

[9] C. Chen, B. He, and X. Tang. Green-aware workload scheduling in geographically distributed data centers. In *Cloud Computing Technology and Science (CloudCom), 2012 IEEE 4th International Conference on*, pages 82–89. IEEE, 2012.

[10] L. Chiaraviglio and I. Matta. An energy-aware distributed approach for content and network management. In *INFOCOM WKSHPS*, pages 337–342. IEEE, 2011.

[11] J. N. Corcoran. Modelling extremal events for insurance and finance. *Journal of the American Statistical Association*, 97(457):360–360, 2002.

[12] P. Cramton. Colombia's forward energy market. 2007.

[13] A. Deoras. Electricity load and price forecasting with matlab, 2010. [Online; accessed 25-April-2015].

[14] A. Gandhi, Y. Chen, D. Gmach, M. Arlitt, and M. Marwah. Minimizing data center SLA violations and power consumption via hybrid resource provisioning. In *Green Computing Conference and Workshops (IGCC), 2011 International*, pages 1–8. IEEE, 2011.

[15] M. Ghamkhari, A. Wierman, and H. Mohsenian-Rad. Energy portfolio optimization of data centers. *IEEE Transactions on Smart Grid*, 2016.

[16] Google. Renewable energy. `http://www.google.com/about/datacenters/renewable/index.html`, 2015. [Online; accessed 25-April-2015].

[17] M. Grant, S. Boyd, and Y. Ye. Cvx: Matlab software for disciplined convex programming, 2008.

[18] Y. Guo, Z. Ding, Y. Fang, and D. Wu. Cutting down electricity cost in internet data centers by using energy storage. In *Proc. IEEE GLOBECOM*, pages 1–5, 2011.

[19] Y. Guo and Y. Fang. Electricity cost saving strategy in data centers by using energy storage. *Parallel and Distributed Systems, IEEE Transactions on*, 24(6):1149–1160, 2013.

[20] V. Gupta, S. Lee, R. Urgaonkar, P. Shenoy, and R. K. Sitaraman. How to cool internet-scale distributed networks on the cheap. 2016.

[21] J. Koomey. Growth in data center electricity use 2005 to 2010. *A report by Analytical Press, completed at the request of The New York Times*, 2011.

[22] H. J. Kushner and G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.

[23] M. Lei, L. Shiyan, J. Chuanwen, L. Hongling, and Z. Yan. A review on the forecasting of wind speed and generated power. *Renewable and Sustainable Energy Reviews*, 13(4):915–920, 2009.

[24] J. Liddle. Amazon found every 100ms of latency cost them 1% in sales. *The GigaSpaces*, 27, 2008.

[25] M. Lin, A. Wierman, L. L. Andrew, and E. Thereska. Dynamic right-sizing for power-proportional data centers. *IEEE/ACM Transactions on Networking (TON)*, 21(5):1378–1391, 2013.

[26] Z. Liu, Y. Chen, C. Bash, A. Wierman, D. Gmach, Z. Wang, M. Marwah, and C. Hyser. Renewable and

cooling aware workload management for sustainable data centers. In *ACM SIGMETRICS Performance Evaluation Review*, volume 40, pages 175–186. ACM, 2012.

[27] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. L. Andrew. Greening geographical load balancing. In *Proc. ACM SIGMETRICS*, pages 233–244, 2011.

[28] Z. Liu, A. Wierman, Y. Chen, B. Razon, and N. Chen. Data center demand response: Avoiding the coincident peak via workload shifting and local generation. *Performance Evaluation*, 70(10):770–791, 2013.

[29] D. Meisner, C. M. Sadler, L. A. Barroso, W.-D. Weber, and T. F. Wenisch. Power management of online data-intensive services. In *Computer Architecture (ISCA), 2011 38th Annual International Symposium on*, pages 319–330. IEEE, 2011.

[30] T. P. Morgan. A rare peek into the massive scale of aws. `http://www.enterprisetech.com/2014/11/14/rare-peek-massive-scale-aws/`, 2014. [Online; accessed 16-May-2016].

[31] E. Nygren, R. Sitaraman, and J. Sun. The Akamai Network: A platform for high-performance Internet applications. *ACM SIGOPS Operating Systems Review*, 44(3):2–19, 2010.

[32] D. S. Palasamudram, R. K. Sitaraman, B. Urgaonkar, and R. Urgaonkar. Using batteries to reduce the power costs of internet-scale distributed networks. In *Proceedings of the Third ACM Symposium on Cloud Computing*, page 11. ACM, 2012.

[33] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs. Cutting the electric bill for internet-scale systems. In *Proc. ACM SIGCOMM*, volume 39, pages 123–134, 2009.

[34] L. Rao, X. Liu, L. Xie, and W. Liu. Minimizing electricity cost: optimization of distributed internet data centers in a multi-electricity-market environment. In *Proc. IEEE INFOCOM*, pages 1–9, 2010.

[35] L. Rao, X. Liu, L. Xie, and Z. Pang. Hedging against uncertainty: A tale of internet data center operations under smart grid environment. *Smart Grid, IEEE Transactions on*, 2(3):555–563, 2011.

[36] R. T. Rockafellar. *Convex analysis*. Princeton university press, 1970.

[37] N. Sharma, P. Sharma, D. Irwin, and P. Shenoy. Predicting solar generation from weather forecasts using machine learning. In *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, pages 528–533. IEEE, 2011.

[38] G. Sinden. Characteristics of the uk wind resource: Long-term patterns and relationship to electricity demand. *Energy Policy*, 35(1):112–127, 2007.

[39] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramaniam. 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*, pages 221–232. ACM, 2011.

[40] J. Whitney and P. Delforge. Data center efficiency assessment. *Issue paper on NRDC (The Natural Resource Defense Council)*, 2014.

[41] L. Yu, T. Jiang, Y. Cao, S. Yang, and Z. Wang. Risk management in internet data center operations under smart grid environment. In *Proc. IEEE SmartGridComm*, pages 384–388, 2012.

[42] Q. Zhang, M. F. Zhani, S. Zhang, Q. Zhu, R. Boutaba, and J. L. Hellerstein. Dynamic energy-aware capacity provisioning for cloud computing environments. In *Proceedings of the 9th international conference on Autonomic computing*, pages 145–154. ACM, 2012.

[43] W. Zheng, K. Ma, and X. Wang. Exploiting thermal energy storage to reduce data center capital and operating expenses. In *Proc. IEEE HPCA*, pages 132–141, 2014.

APPENDIX

A. PROOFS FOR SECTION 3

A.1 Proof of Theorem 1

To prove Theorem 1, we first show that the real-time objective is jointly convex with respect to $(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda})$.

LEMMA 5. *$\tilde{F}^r$ as defined in* (4) *is jointly convex with respect to $(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda})$ over $\mathbb{R}_+^N \times C(L^r)$.*

PROOF. We rewrite

$$\tilde{F}^r(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda}, \boldsymbol{p}^r, \boldsymbol{w}^r) = \sum_{i=1}^{N} p_i^l [m_i - w_i^r - q_i^l]_+ + \sum_{i=1}^{N}\sum_{j=1}^{J} \lambda_{ij} h_{ij}(m_i, \lambda_i). \quad (9)$$

Since $m_i - w_i^r - q_i^l$ is an affine function, and $[\cdot]_+$ is convex and non-decreasing, $\sum_{i=1}^N p_i^l [m_i - w_i^r - q_i^l]_+$ is jointly convex with respect to $(\boldsymbol{q}^l, \boldsymbol{m})$.

Since $\lambda_{ij} h_{ij}(m_i, \lambda_i)$ is jointly convex with respect to $(\boldsymbol{m}, \boldsymbol{\lambda})$, $\tilde{F}^r$ is jointly convex with respect to $(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda})$ because the summation of convex functions are convex. □

PROOF OF THEOREM 1. From Lemma 5, we know that the real time objective function $\tilde{F}^r(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda}, \boldsymbol{p}^r, \boldsymbol{w}^r)$ is jointly convex with respect to $(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda})$. It then follows that

$$F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r) = \min_{(\boldsymbol{m}, \boldsymbol{\lambda}) \in C(L^r)} \tilde{F}^r(\boldsymbol{q}^l, \boldsymbol{m}, \boldsymbol{\lambda}, \boldsymbol{p}^r, \boldsymbol{w}^r)$$

is convex with respect to $\boldsymbol{q}^l$ (see [7]). Finally, since the expectation operation preserves convexity, we conclude that $F^l(\boldsymbol{q}^l)$ is convex with respect to $\boldsymbol{q}^l$. □

A.2 Proof of Theorem 2

This section is devoted to the proof of Theorem 2. To prove Theorem 2, it suffices to show that

$$\begin{aligned} \frac{\partial \mathbb{E}\left[F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)\right]}{\partial q_i^l} &= \mathbb{E}\left[\frac{\partial F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)}{\partial q_i^l}\right] \\ &= -\mathbb{E}\left[\varrho_i(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)\right]. \end{aligned} \quad (10)$$

The first step is to prove that the Lagrange multiplier of GLB-RT corresponding to the constraint (3f) is unique.

LEMMA 6. *With probability 1, GLB-RT has a unique Lagrange multiplier, denoted $\varrho_i(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$, corresponding to the constraint* (3f).

PROOF. In this proof, for notational simplicity, we suppress the dependence of the primal and dual solutions of GLB-RT on $(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$. Consider a primal solution of GLB-RT $(\boldsymbol{q}^r, \boldsymbol{m}, \boldsymbol{\lambda})$ with $\boldsymbol{m} > 0$. Such a solution exists with probability 1, since $\boldsymbol{w}^r > 0$ with probability 1.

Now any dual solution must satisfy the KKT conditions. This implies the following conditions. (Since the constraint $\lambda_i \leq m_i\mu_i$ is never binding, the corresponding Lagrange multiplier $\sigma_i = 0$ and does not feature in the following.)

$$\frac{\partial}{\partial m_i}\sum_{i=1}^{N}\sum_{j=1}^{J}\lambda_{ij}h_{ij}(m_i,\lambda_i) + \bar{\omega}_i - \underline{\omega}_i + \varrho_i = 0 \tag{11}$$

$$\bar{\omega}_i(m_i - M_i) = 0;\ \bar{\omega}_i \geq 0, m_i \leq M_i \tag{12}$$

$$\underline{\omega}_i m_i = 0;\ \underline{\omega}_i \geq 0, m_i \geq 0 \tag{13}$$

$$p_i^r - \kappa_i - \varrho_i = 0 \tag{14}$$

$$\kappa_i q_i^r = 0; \kappa_i \geq 0, q_i^r \geq 0 \tag{15}$$

$$\varrho_i(-q_i^r + m_i - w_i^r - q_i^l) = 0; \tag{16}$$

$$\varrho_i \geq 0, q_i^r \geq m_i - w_i^r - q_i^l \tag{17}$$

We now argue that ϱ_i is unique for all i. Consider the following two cases.

Case 1: $w_i^r + q_i^l > M_i$. In this case, it follows that $m_i < w_i^r + q_i^l + q_i^r$. Using (16), we conclude that $\varrho_i = 0$.

Case 2: $w_i^r + q_i^l < M_i$. Here we consider two sub-cases.

Case 2a: $m_i = M_i$. In this case, it follows that $q_i^r > 0$, which implies that $\kappa_i = 0$ (by (15)). Thus, we have, using (14), that $\varrho_i = p_i^r$.

Case 2b: $m_i < M_i$. In this case, since $m_i \in (0, M_i)$, we have $\bar{\omega}_i = \underline{\omega}_i = 0$ (by (12) and (13)). Thus, from (11), we have

$$\varrho_i = -\frac{\partial}{\partial m_i}\sum_{i=1}^{N}\sum_{j=1}^{J}\lambda_{ij}h_{ij}(m_i,\lambda_i). \tag{18}$$

Since the event $w_i^r + q_i^l = M_i$ has zero probability, we may ignore this case. This completes the proof. □

Given Lemma 6, it follows from standard sensitivity analysis in convex optimization (see Section 6.5.3 and 6.5.4 in [6]) that

$$\frac{\partial F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)}{\partial q_i^l} = -\varrho_i(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r). \tag{19}$$

This proves the second equality in (10). Thus, to complete the proof of Theorem 2, it only remains to justify the interchange of the partial derivative and the expectation in the first equality. We justify this interchange by invoking the dominated convergence theorem as follows.

Let e_i denote a column vector in $\mathbb{R}^N$, with 1 in the ith entry and 0 elsewhere.

LEMMA 7. *For any $\delta \neq 0$ and $i \in N$,*

$$\left|\frac{F^{*r}(\boldsymbol{q}^l + \delta e_i, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r) - F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)}{\delta}\right| \leq p_i^r.$$

PROOF. It is easy to see that

$$F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r) \leq \delta p_i^r + F^{*r}(\boldsymbol{q}^l + \delta e_i, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r).$$

The statement of Lemma 7 follows from the fact that the function $F^{*r}(\boldsymbol{q}^l + \delta e_i, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)$ is non-increasing with respect to δ. □

Since $\mathbb{E}[p_i^r] < \infty$, it follows from the dominated convergence theorem that

$$\mathbb{E}\left[\lim_{\delta\to 0}\frac{F^{*r}(\boldsymbol{q}^l + \delta e_i, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r) - F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)}{\delta}\right]$$
$$= \lim_{\delta\to 0}\frac{\mathbb{E}\left[F^{*r}(\boldsymbol{q}^l + \delta e_i, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)\right] - \mathbb{E}\left[F^{*r}(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r)\right]}{\delta}.$$

This proves the first equality in (10), and completes the proof of the theorem.

A.3 Proof of Lemma 3

We assume there an optimal solution S such that $\lambda_i > 0$ and $0 < m_i < w_i^r + q_i^l$. $\lambda_i = 0$ or $m_i = 0$ is not ignored because it is equivalent to shutting down data center i. Here, $\underline{\omega}_i = \varrho_i = 0$.

If $w_i^r + q_i^l < M_i$ and $m_i < w_i^r + q_i^l$, $\bar{\omega}_i = 0$. $\bar{\omega}_i = \underline{\omega}_i = \varrho_i = 0$ results in λ_i/m_i becomes zero as (11). $\lambda_i = 0$ or $m_i = \infty$ contradicts the assumption. So, $m_i \geq w_i^r + q_i^l$ if $w_i^r + q_i^l < M_i$.

If $w_i^r + q_i^l \geq M_i$, we assume that $m_i < M_i$. $\bar{\omega}_i = \underline{\omega}_i = \varrho_i = 0$ again leads to the contradiction to the assumption. So, $m_i = M_i$ if $w_i^r + q_i^l \geq M_i$.

B. CONVERGENCE OF SGA

This section is devoted to the proof of Theorem 4. Invoking Theorem 2.1 in [22], the almost sure convergence of the iterates of SGA to the set of optimal solutions of EP-LT holds if the following two conditions are satisfied.

1. $\nabla F^l : \mathbb{R}_+^N \to \mathbb{R}^N$ is continuous.
2. $\sup_{q^l \in \mathbb{R}_+^N} \mathbb{E}\left[(\varrho_i(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r))^2\right] < \infty$.

Condition (1) above holds since the gradient of a differentiable convex function is continuous; see Theorem 25.5 in [36]. Condition (2) holds since

$$\varrho_i(\boldsymbol{q}^l, \boldsymbol{p}^r, \boldsymbol{L}^r, \boldsymbol{w}^r) \leq p_i^r$$

(see (19) and Lemma 7). This completes the proof.