Electricity Cost Saving Strategy in Data Centers by Using Energy Storage

Yuanxiong Guo, *Student Member, IEEE*, and Yuguang Fang, *Fellow, IEEE*

Abstract—Electricity expenditure comprises a significant fraction of the total operating cost in data centers. Hence, cloud service providers are required to reduce electricity cost as much as possible. In this paper, we consider utilizing existing energy storage capabilities in data centers to reduce electricity cost under wholesale electricity markets, where the electricity price exhibits both temporal and spatial variations. A stochastic program is formulated by integrating the center-level load balancing, the server-level configuration, and the battery management while at the same time guaranteeing the quality-of-service experience by end users. We use the Lyapunov optimization technique to design an online algorithm that achieves an explicit tradeoff between cost saving and energy storage capacity. We demonstrate the effectiveness of our proposed algorithm through extensive numerical evaluations based on real-world workload and electricity price data sets. As far as we know, our work is the first to explore the problem of electricity cost saving using energy storage in multiple data centers by considering both the spatial and temporal variations in wholesale electricity prices and workload arrival processes.

Index Terms—Cloud computing, electricity cost, data center, energy storage, Lyapunov optimization, wholesale electricity market

1 Introduction

With the popularity of cloud computing [2], more and more data centers are envisioned to be built in the future in order to meet the growing demand for large-scale computing resources. It is common for a cloud service provider to have multiple data centers each having hundreds of thousands of servers. Those data centers are geographically distributed for reliability as well as performance improvement. A critical issue in the operations of those data centers is the energy consumption, including both servers and air conditioners. According to the estimation from [3], large companies such as Google and Microsoft pay tens of millions of dollars for just electricity usage every year, and 30-50 percent percentage of operational expenses in data centers come from electricity. Therefore, minimizing the electricity cost is receiving more and more attention. However, saving electricity cost and improving performance are usually in conflict with each other; thus, joint optimization is needed.

A natural way to reduce electricity cost is to conserve energy consumption or to improve the energy efficiency such that the same amount of workload can be served with less energy. Note that in large-scale data centers, computing equipments generally exhibit high power intensity with all of its consumed electric power converted to heat. In order to ensure the reliable operation of data centers, air conditioning is required to extract the heat dissipated by the IT computing devices. Thus, additional power is required to operate the cooling system. Power Usage Effectiveness (PUE), which measures the ratio of total building power to IT power, i.e., the power consumed by the actual computing equipment, is used to judge the energy efficiency of a data center. It is reported that PUE is nearly 2 for typical data centers [4]. Various engineering techniques such as advanced cooling, virtualization, direct current (DC) power, multicore servers, etc., have been employed to improve the PUE.

On the other hand, more and more electricity markets are undergoing deregulation where the electricity market operators offer dynamic electricity rates to large industrial and commercial customers instead of traditional flat rates at the retail level. Therefore, minimization of electricity consumption does not necessarily translate into that of the electricity cost since the cost should be the price times the energy amount. Geographical load balancing [3], [5] has been proposed to utilize the variation of electricity prices in wholesale electricity markets so as to provide significant cost savings for data centers. The basic idea is to route more traffic to data centers where the electricity price is lower. Although those techniques are effective in practice, a largely ignored factor is the existence of energy storage facilities within data centers, which can provide further cost saving if utilized intelligently in combination of the previous techniques. Comparing with existing techniques for power cost reduction, the method of energy storage has no performance degradation.

Data centers have uninterrupted power supply (UPS) units to keep them powered using stored energy in case of electric utility failure, which is their primary power source, before the backup diesel generation can start up and provide power as secondary power source. Usually, the transition to use diesel generation takes only 10-20 seconds while UPS units have enough capacity to power the data center at its

- *Y. Guo is with the Department of Electrical and Computer Engineering, University of Florida, 446 New Engineering Building, Gainesville, FL 32603. E-mail: guoyuanxiong@ufl.edu.*
- *Y. Fang is with the Department of Electrical and Computer Engineering, University of Florida, 435 Engineering Building, PO Box 116130, Gainesville, FL 32611. E-mail: fang@ece.ufl.edu.*

Manuscript received 29 Feb. 2012; revised 10 June 2012; accepted 15 June 2012; published online 22 June 2012.
Recommended for acceptance by V.B. Misic, R. Buyya, D. Milojicic, and Y. Cui.
For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDSSI-2012-02-0176.
Digital Object Identifier no. 10.1109/TPDS.2012.201.

1045-9219/13/$31.00 © 2013 IEEE Published by the IEEE Computer Society

maximum power need between 5-30 minutes. This excess energy storage capacity can be used to save the electricity cost by the simple intuition of charging when the electricity price is low while discharging when the electricity price is high in the utility grid.

In this paper, we investigate the problem of exploiting the UPS units within data centers to minimize the cloud service provider's electricity cost. We propose a joint load balancing, server configuration, and battery management scheme for multiple distributed data centers. Since the traffic arrivals and electricity prices are both random processes with possibly unknown statistics, the problem is formulated as a stochastic program and then, an efficient online algorithm based on the Lyapunov optimization technique [6] is proposed to solve it.

Our contribution can be summarized as follows:

- We investigate the problem of minimizing the total electricity cost of multiple data centers for a cloud service provider under wholesale electricity markets by taking into account the batteries within these data centers.
- We formulate the problem as a stochastic program, which captures the center-level load balancing, the server-level configuration, and the battery management while at the same time guaranteeing the quality-of-service (QoS) experience by end users.
- We propose an efficient online algorithm based on the Lyapunov optimization technique to obtain the optimal joint load balancing, server configuration, and battery management scheme for the total electricity cost minimization. Moreover, our algorithm offers an explicit tradeoff between the cost saving and the battery capacity.
- We evaluate our algorithm based on real-world data sets and the results show that our approach can achieve significant electricity cost saving.

The rest of this paper is organized as follows. Section 2 reviews some related studies in electricity cost reduction in data centers. Section 3 presents the models of electricity cost in multiple data centers, which is formulated as a stochastic program to minimize the time-average expected electricity cost. Section 4 solves the optimization problem by first considering a relaxed problem and then, using the Lyapunov optimization technique to design a control algorithm to approximately solve the original problem. Section 5 gives the algorithmic performance analysis and Section 6 gives the numerical evaluation results based on real-world data sets. Finally, Section 7 concludes the paper.

2 Related Work

The huge energy consumption in data centers has motivated a lot of research to reduce the electricity cost in data centers. These studies can be roughly divided into two categories: one is from the perspective of hardware design and engineering; the other is from the perspective of algorithmic design.

In the first category, energy-efficient chips, multicore servers, DC power supplies, advanced cooling systems, and virtulization have been developed [7]. These techniques have been used to improve the PUE for data centers (See [8] for a survey on these issues).

In the second category, the research can be divided into the following three different levels. The first level is the server level, where only the power consumption of a single server is considered. A widely used technique is the dynamic voltage/frequency scaling (DVFS), where the operating voltage and frequency of the server's CPU can be adjusted according to the intensity of the workload on the server. Since the first analytical study of DVFS by Yao et al. [9], the scheduling and speed scaling algorithms to minimize the total energy used in order to meet job deadlines have been addressed in [10]. The objective of minimizing the average response time given an energy budget is addressed in [11], while the objective of minimizing a weighted combination of expected response time and energy usage per job is considered in [12].

The second level is the data center level. Dynamic cluster server configuration (DCSC) has been proposed to optimally adjust the number of active servers in data centers while satisfying the QoS requirement for electricity cost reduction. Lin et al. propose a novel online algorithm for cost reduction to dynamically right size a data center, which is proven to be 3-competitive, while taking into account the switching cost during turning on/off servers [13].

The third level is the interdata center level, which is based on the observation that the electricity price is different across different time and locations under wholesale electricity markets. Qureshi et al. are the first to discuss the opportunity of utilizing such electricity price variation to reduce total electricity cost by distributing more traffic to data centers with low electricity price [3]. Rao et al. investigate the problem of minimizing the total electricity cost for data centers in a multielectricity-market environment subject to QoS guarantee and propose a linear programming formulation to approximately solve it [5]. These studies focus on directly reducing the total electricity cost by exploiting the spatial variation of electricity prices. However, none of the aforementioned work considers using available energy storage capabilities, typically UPS units, in data centers to further reduce the electricity cost.

Our work is mainly motivated by Urgaonkar et al. [14], which considers the case of a single data center with energy storage for time-varying electricity price under a wholesale electricity market. The system implementation issue of using UPS units to help reduce electricity cost is analyzed in [15]. Different from previous studies, our work considers the total electricity cost minimization of a cloud service provider having multiple data centers with energy storage under both time-varying and location-varying electricity prices. In our preliminary work [1], we assume that the server consumes a fixed amount of power when turned on and the battery cost is not considered. However, in this paper, we take into account the battery cost as well as realistic server power consumption model. Moreover, extensive numerical evaluation results based on real-world data sets are included in this paper.

3 Model and Formulation

We now describe the models we use in this paper to minimize the time-average expected electricity cost in data centers. Assume the system is discrete time with time period

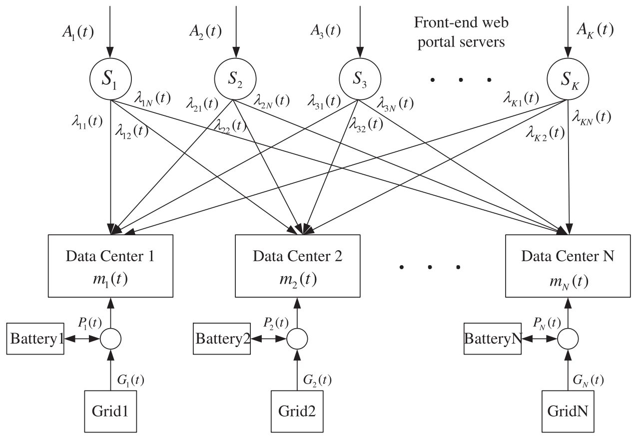


Fig. 1. Block diagram for the system model.

matching the timescale at which traffic distribution, server configuration, and charging/discharging decisions can be updated (e.g., 10 min).

We consider a cloud service provider having N geographically distributed data centers, denoted by $\mathcal{D} = \{D_1, \ldots, D_N\}$ and K front-end proxy servers, denoted by $\mathcal{S} = \{S_1, \ldots, S_K\}$. Each data center D_i has a total number of M_i homogeneous servers. The system operates in slotted period, i.e., $t = \{0, 1, \ldots\}$. The block diagram of our system model is shown in Fig. 1, which is described in detail as follows.

3.1 The Workload Model

In every period t, customer requests arrive at each front-end proxy server. We denote the average arrival rate of workload at S_j by $A_j(t)$, $j \in \{1, \ldots, K\}$, where $\mathbf{A}(t) = (A_1(t), \ldots, A_K(t))$ denotes the traffic arrival vector. The workload arrival rate distributed from the front-end proxy server S_j to the data center D_i is denoted as $\lambda_{ji}(t)$. This can be done by dynamically generated DNS responses, HTTP redirections, or using persistent HTTP proxies to tunnel requests. We assume that there exists a proxy/DNS server colocated with each request source. Therefore, we have

$$\sum_{i=1}^{N} \lambda_{ji}(t) = A_j(t), \quad \forall j = 1, \ldots, K. \tag{1}$$

$$\lambda_{ji}(t) \geq 0. \tag{2}$$

Define the total arrival rate distributed to data center D_i as $\lambda_i(t)$ and the distributed workload vector as $\boldsymbol{\lambda}(t) = (\lambda_1(t), \ldots, \lambda_N(t))$. Then, we have

$$\lambda_i(t) = \sum_{j=1}^{K} \lambda_{ji}(t), \quad \forall i = 1, \ldots, N. \tag{3}$$

3.2 The Battery Model

We assume that each data center possesses some kind of battery. For each data center D_i, we denote by $E_{i,max}$ the battery capacity, by $E_i(t)$ the energy level of the battery at period t, and by $P_i(t)$ the power (energy per period) charged to (when $P_i(t) > 0$) or discharged from (when $P_i(t) < 0$) the battery during period t. Assume that the battery energy leakage is negligible and batteries at data centers operate independently of each other. Then, we model the dynamics of the battery energy level by

$$E_i(t+1) = E_i(t) + P_i(t). \tag{4}$$

For each data center D_i, the battery usually has an upper bound on the charge rate, denoted by $P_{i,max}$, and an upper bound on the discharge rate, denoted by $P_{i,min}$. $P_{i,max}$ and $P_{i,min}$ are positive constants depending on the physical property of the battery. Therefore, we have the following constraint on $P_i(t)$:

$$-P_{i,min} \leq P_i(t) \leq P_{i,max}. \tag{5}$$

The battery energy level should be always nonnegative and cannot exceed the battery capacity. Therefore, in each time period t, we need to ensure that for each data center D_i,

$$0 \leq E_i(t) \leq E_{i,max}. \tag{6}$$

From constraints (4), (5), and (6), we get the following equivalent constraints in each period t for data center D_i:

$$P_i(t) \geq -\min\{P_{i,min}, E_i(t)\}, \tag{7}$$

$$P_i(t) \leq \min\{P_{i,max}, E_{i,max} - E_i(t)\}. \tag{8}$$

However, the cost of using battery cannot be ignored. In practice, there are limited times of charging/discharing cycles for each battery. Besides, conversion loss occurs both in charging and discharging processes. Stored energy is also subject to leakage with time. All these factors depend on how fast/much/often it is charged and discharged. Instead of modeling these factors exactly, we use an amortized cost C_b (in unit of dollars) to model the impact of per charging or discharging operation on the battery during one period. Therefore, during one time period, an operating cost of C_b is incurred whenever the battery is charging ($P_i(t) > 0$) or discharging ($P_i(t) < 0$).

3.3 The QoS Model

In practice, according to the service level agreement (SLA) between the service provider and customers, customer requests should have some kind of QoS requirements. In this paper, we use the average response time as the QoS metric. As in [5], we use an $M/M/n$ queuing model to analyze the average response time in data center D_i when the traffic arrival rate is $\lambda_i(t)$ and there are $m_i(t)$ active servers, each with service rate $\mu_i(t)$. Note that $m_i(t)$ is an integral variable and has the maximum value M_i at each data center D_i. Also, there exists the maximum service rate $\mu_{i,max}$ for each server in data center D_i. When there is traffic distributed into D_i, using the results from queuing theory [16], the average response time $W_i(t)$ is $\frac{1}{m_i(t)\mu_i(t)-\lambda_i(t)} P_Q$ where P_Q is the queuing probability. Without loss of generality in a data center, we assume the servers are always busy if turned on. Hence, $P_Q = 1$ and $W_i(t) = \frac{1}{m_i(t)\mu_i(t)-\lambda_i(t)}$. To meet the QoS requirement of customers, the maximum average response time $W_{i,max}$ is imposed on each data center D_i. Therefore, we have the following QoS constraints:

$$m_i(t)\mu_i(t) - \lambda_i(t) \geq \frac{1}{W_{i,max}}, \quad \forall i = 1, \ldots, N, \tag{9}$$

$$0 \le m_i(t) \le M_i, \quad m_i \in \mathbb{N}, \quad \forall i = 1, \dots, N, \tag{10}$$

$$0 \le \mu_i(t) \le \mu_{i,max}, \quad \forall i = 1, \dots, N. \tag{11}$$

3.4 The Power Consumption Model

In each time period t, the normal power consumption $H_i(t)$, including the cooling energy consumption at each data center D_i, by running $m_i(t)$ servers at rate $\mu_i(t)$ can be approximated by the following formula [17]:

$$H_i(t) = m_i(t) \cdot (\alpha_i \mu_i^{\gamma_i}(t) + \beta_i) \cdot \mathrm{PUE}_i, \tag{12}$$

where α_i, β_i, γ_i, and PUE_i are constants determined by the data center D_i. Specifically, β_i is the average idle power consumption of a server, and $\alpha_i \mu_i^{\gamma_i}(t) + \beta_i$ gives the power consumption of one server running at rate $\mu_i(t)$ at D_i. PUE_i is the ratio of the total building power (including cooling power) to IT server power, whose value lies between 1.3 and 2 in today's energy-efficient data centers [4].

Due to the introduction of energy storage, the total amount of energy $G_i(t)$ drawn from the grid to supply the data center D_i during time period t is given by

$$G_i(t) = H_i(t) + P_i(t), \quad \forall i = 1, \dots, N. \tag{13}$$

Specifically, when $P_i(t) > 0$, some energy drawn from the grid is used to charge the battery besides serving the normal data center operation. When $P_i(t) < 0$, some energy is discharged from the battery to supplement the energy drawn from the grid so as to meet the energy demand of the data center.

3.5 The Electricity Price Model

The electric power grid in US is organized into different reliability regions, where each region has its own regional transmission organization (RTO) or independent system operator (ISO) [18]. The RTO or ISO is a central authority that directs the flow of electricity between generators and consumers and ensures the reliability of the grid. It also operates wholesale electricity markets, which usually include day-ahead and real-time electricity markets. The electricity prices in these markets are determined by the clearing processes of supply and demand bids while satisfying the transmission constraints.

As analyzed in [3], the electricity prices in wholesale electricity markets have both spatial and temporal variations. At each data center D_i, we assume a time-varying electricity price $C_i(t)$ with the maximum value $C_{i,max}$ and the minimum value $C_{i,min}$, respectively. Denote $\mathbf{C}(t) = (C_1(t), \dots, C_N(t))$ as the electricity price vector and $\mathbf{G}(t) = (G_1(t), \dots, G_N(t))$ as the grid energy consumption vector. We further assume that $\mathbf{C}(t)$ and $\mathbf{G}(t)$ are independent. Different data centers may have different electricity prices at the same time due to being located in different electricity markets.

3.6 The Cost Minimization Problem with Energy Storage

As discussed before, the total electricity cost of N data centers during time period t is given by the following:

$$\sum_{i=1}^{N} \{G_i(t)C_i(t) + \mathbf{1}_{\{P_i(t)\neq 0\}} C_b\}. \tag{14}$$

In this paper, we are interested in choosing the following three control decisions to minimize the long-term time-average expected electricity cost: 1) the workload distributed from the front-end web portals to different data centers—$\boldsymbol{\lambda}(t)$; 2) the number of active servers at different data centers—$\mathbf{m}(t)$ and the corresponding service rates—$\boldsymbol{\mu}(t)$; 3) the charge/discharge rate at different data centers—$\mathbf{P}(t)$. Based on the models above, our problem can be formulated as the following stochastic program, called **P1**:

$$\min \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \mathbb{E}\{G_i(t)C_i(t) + \mathbf{1}_{\{P_i(t)\neq 0\}} C_b\},$$

subject to constraints (1), (2), (6), (7), (8), (9), (10), (11), and (13), where the constraints are for each time period t and data center D_i.

4 Proposed Solution

One challenge of solving the stochastic optimization problem above is the unawareness of future workload arrivals as well as time-varying and location-varying electricity prices. Moreover, the constraints on $E_i(t)$ bring the "time-coupling" property to the stochastic optimization problem above. It means that the current control action may impact the future control actions, making it more challenging to solve. As mentioned before, the statistics of $\mathbf{A}(t)$ and $\mathbf{C}(t)$ may not be known and we need to design an optimal control algorithm under uncertainty. We use the recently developed technique of Lyapunov optimization [6]. The algorithm we propose can achieve the range of $O(1/V)$ within the optimal objective value, where V is a parameter related to the battery capacity of each data center D_i. One salient feature of our algorithm is that it does not need any future knowledge of the system and can be easily implemented online.

4.1 Relaxed Problem

Before giving the solution to our original problem **P1**, we consider a relaxed problem. Define the time-average expected value of charge or discharge rate at data center D_i under any feasible control policy of **P1** as follows:

$$\overline{P_i} = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{P_i(t)\}. \tag{15}$$

Since the battery energy level is evolving according to (4), summing over all $t \in \{0, 1, 2, \dots, T-1\}$, and taking expectation of both sides, we have

$$\mathbb{E}\{E_i(T)\} - E_{i,ini} = \sum_{t=0}^{T-1} \mathbb{E}\{P_i(t)\},$$

where $E_{i,ini} = E_i(0)$ is the initial battery energy level at data center D_i. As $0 \le E_i(t) \le E_{i,max}$ for all time periods t, dividing both sides by T, and taking $T \to \infty$ yields $\overline{P_i} = 0$. Hence, we have the following relaxed problem, called **P2**:

$$\min \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \mathbb{E}\{G_i(t)C_i(t) + \mathbf{1}_{\{P_i(t)\neq 0\}} C_b\},$$

subject to constraints (1), (2), (5), (9), (10), (11), (13), and

$$\overline{P_i} = 0,$$

where the constraints are for each time period t and data center D_i.

Denote the optimal objective value of **P1** as Q^{OPT} and the optimal objective value of **P2** as Q^{REL}. As discussed before, any feasible solution to **P1** is also a feasible solution to **P2**. Hence, $Q^{REL} \leq Q^{OPT}$. Note that **P2** is a decoupled control problem since no correlation exists in any constraint. From the framework of Lyapunov optimization [6], we have the following theorem for the solution to **P2**:

Theorem 1. *If $\mathbf{C}(t)$ and $\mathbf{A}(t)$ are i.i.d. over slots, then there exists a stationary, randomized policy that takes control decisions $\boldsymbol{\lambda}^{stat}(t)$, $\mathbf{m}^{stat}(t)$, $\boldsymbol{\mu}^{stat}(t)$, and $\mathbf{P}^{stat}(t)$ every period purely as a function (possibly randomized) of the current workload vector $\mathbf{A}(t)$ and the electricity price vector $\mathbf{C}(t)$ while satisfying the constraints of* **P2** *and providing the following guarantees:*

$$\mathbb{E}\{P_i^{stat}(t)\} = 0,$$

$$\mathbb{E}\left\{\sum_{i=1}^{N}\left[G_i^{stat}(t)C_i(t) + \mathbf{1}_{\{P_i^{stat}(t)\neq 0\}}C_b\right]\right\} = Q^{REL},$$

where the expectations above are with respect to the stationary distributions of $\mathbf{A}(t)$, $\mathbf{C}(t)$, and the randomized control decisions.

Proof. It can be proven using Caratheodory's theorem in [6] and is similar to that in [14]. It is omitted here for brevity. □

In order to derive such a policy, we need to know the statistical distributions of all combinations of $\mathbf{C}(t)$ and $\mathbf{A}(t)$, which usually has the problem of "curse of dimensionality" [19] if solved by dynamic programming. Moreover, this control policy may not be a feasible solution to **P1**. Instead, we use the existence of such a policy to help us design our control policy that meets all constraints of **P1** and derive the algorithmic performance of our algorithm as illustrated in the proof of our algorithmic properties later.

4.2 Our Proposed Algorithm

The idea of our algorithm is to construct a Lyapunov scheduling algorithm with perturbed weights for determining the traffic distribution, data center sizing, service rate, and charging/discharging decisions. By carefully perturbing the weights, we can ensure that whenever we charge or discharge the battery, the energy level in the battery always lies in the feasible region.

First, we define a modified Lyapunov function as follows:

$$L(t) \triangleq \frac{1}{2}\sum_{i=1}^{N}[E_i(t) - VC_{i,max} - P_{i,min}]^2. \tag{16}$$

To simplify the notation, we define a variable $S_i(t)$ for each data center D_i, $i \in \{1, \ldots, N\}$ as follows:

$$S_i(t) = E_i(t) - VC_{i,max} - P_{i,min}. \tag{17}$$

Let $\mathbf{S}(t) = (S_1(t), \ldots, S_N(t))$. It is obvious that $S_i(t)$ is just a shifted version of $E_i(t)$ and has the same dynamics as $E_i(t)$ with the following equation:

$$S_i(t+1) = S_i(t) + P_i(t). \tag{18}$$

Then, the Lyapunov function can be rewritten as follows:

$$L(t) \triangleq \frac{1}{2}\sum_{i=1}^{N} S_i^2(t).$$

Now, define the one-period conditional Lyapunov drift as follows:

$$\triangle(t) = \mathbb{E}\{L(t+1) - L(t)|\mathbf{S}(t)\}. \tag{19}$$

Here, the expectation is taken over the randomness of electricity prices and workload arrivals, as well as the randomness in choosing the control actions. Then, following the Lyapunov optimization framework, we add a function of the expected electricity cost over one period (i.e., the penalty function) to (19) to obtain the following *drift-plus-penalty* term:

$$\triangle_V(t) \triangleq \triangle(t) + V\mathbb{E}\left\{\sum_{i=1}^{N}[G_i(t)C_i(t) + \mathbf{1}_{\{P_i(t)\neq 0\}}C_b]|\mathbf{S}(t)\right\}. \tag{20}$$

We have the following lemma regarding the *drift-plus-penalty* term:

Lemma 1. *Under any feasible action that can be implemented at period t, we have*

$$\begin{aligned}\triangle_V(t) \leq B &+ \sum_{i=1}^{N}\mathbb{E}\{S_i(t)P_i(t) \mid \mathbf{S}(t)\}\\ &+ V\sum_{i=1}^{N}\mathbb{E}\{G_i(t)C_i(t) + \mathbf{1}_{\{P_i(t)\neq 0\}}C_b \mid \mathbf{S}(t)\},\end{aligned} \tag{21}$$

where $B \triangleq \sum_{i=1}^{N}\frac{\max\{P_{i,max}^2, P_{i,min}^2\}}{2}$.

Proof. From (18), squaring both sides, we have for each data center D_i,

$$\frac{S_i^2(t+1) - S_i^2(t)}{2} = \frac{P_i^2(t)}{2} + S_i(t)P_i(t). \tag{22}$$

Moreover, we have the following inequality:

$$\frac{P_i^2(t)}{2} \leq \frac{\max\{P_{i,max}^2, P_{i,min}^2\}}{2} \triangleq B_i.$$

Taking expectations of both sides of (22) given $\mathbf{S}(t)$ and summing over all data centers D_i, we have

$$\triangle(t) \leq \sum_{i=1}^{N}B_i + \sum_{i=1}^{N}\mathbb{E}\{P_i(t)S_i(t) \mid \mathbf{S}(t)\}.$$

Adding penalty term

$$V\sum_{i=1}^{N}\mathbb{E}\{[G_i(t)C_i(t) + \mathbf{1}_{\{P_i(t)\neq 0\}}C_b] \mid \mathbf{S}(t)\}$$

to both sides of the inequality above, we arrive at the conclusion. □

We now present our algorithm. The main design principle of our algorithm is to choose control actions that minimize the R.H.S. of (21) subject to the constraints in **P2**. Our algorithm works as follows.

Algorithm 1: Cost Minimization with Energy Storage

```
foreach Time period t do
    Observe the system states A(t), C(t), and S(t);
    Choose control decisions λ*(t), m*(t), μ*(t), and
    P*(t) as the optimal solution to the following
    optimization problem, called P3:
    Minimize
        Σ_{i=1}^{N} { S_i(t)P_i(t) + V [G_i(t)C_i(t) + 1_{P_i(t)≠0} C_b] },
    subject to constraints (1), (2), (5), (9), (10), (11),
    and (13) where the constraints are for each time
    period t and data center D_i;
    Update E_i(t) according to the dynamics (4);
```

$$\sum_{i=1}^{N}\Big\{S_i(t)P_i(t)+V\left[G_i(t)C_i(t)+\mathbf{1}_{\{P_i(t)\neq 0\}}C_b\right]\Big\},$$

Note that the algorithm above only requires the knowledge of the instant values of electricity prices $\mathbf{C}(t)$, traffic arrival rates $\mathbf{A}(t)$, and battery energy levels $\mathbf{E}(t)$. It does not require any knowledge of the statistics of these stochastic processes. The remaining challenge is to solve **P3**, which is discussed below.

4.3 Solution to P3

As we can observed, for each time period t, the optimization problem above is a mixed-integer nonlinear programming (MINLP), which is NP-hard in general. As in [20], by using the KKT conditions, we can obtain that the optimal solution, if exists, must satisfy the following:

$$\lambda_i^*(t) = m_i^*(t)\mu_i^*(t) - \frac{1}{W_{i,max}}. \tag{23}$$

It can be observed that once the optimal $m_i^*(t)$ and $\mu_i^*(t)$ are obtained, we can solve the corresponding $\lambda_{ij}^*(t)$ by the following two equations:

$$\sum_{i=1}^{N}\lambda_{ji}^*(t) = A_j(t), \quad \lambda_{ji}^*(t)\geq 0, \tag{24}$$

$$\sum_{j=1}^{K}\lambda_{ji}^*(t) = m_i^*(t)\mu_i^*(t) - \frac{1}{W_{i,max}}. \tag{25}$$

In the following part, we focus on how to obtain the optimal $m_i^*(t)$ and $\mu_i^*(t)$.

We first define $I_i(t)$ as the indicator variable to describe the battery usage associated with each data center D_i during time period t. When the battery is used (either charging or discharging) at data center D_i, $I_i(t)=1$. Otherwise, $I_i(t)=0$. Then, the optimization problem can rewritten as follows, named **P4**:

Minimize

$$\sum_{i=1}^{N}\{(S_i(t)+VC_i(t))P_i(t)+VC_bI_i(t)+VC_i(t)H_i(t)\},$$

s.t.

$$\sum_{i=1}^{N}m_i(t)\mu_i(t) = \sum_{j=1}^{K}A_j(t)+\sum_{i=1}^{N}\frac{1}{W_{i,max}},$$

$$0\leq m_i(t)\leq M_i, \quad m_i(t)\in\mathbb{N},$$

$$0\leq \mu_i(t)\leq \mu_{i,max},$$

$$-P_{i,min}I_i(t)\leq P_i(t)\leq P_{i,max}I_i(t),$$

$$I_i(t)\in\{0,1\},$$

$$0\leq P_i(t)+H_i(t)\leq G_{i,max},$$

where the constraints are for each time period t and data center D_i.

A general MINLP is known to be NP-hard and no efficient solutions exist when the problem size is large because the search space would increase exponentially. However, in some practical situations, the MINLP problem often have some special structures that can be exploited for designing effective solutions. One particular situation, as in our problem, is that by fixing the discrete variables first, the remaining problem becomes convex for continuous variables. In this paper, we use the technique of generalized Benders decomposition [21] to solve it. The proposed algorithm is expected to converge to the optimal solution within a finite number of iterations. To simplify the notation, we ignore the time t in the following algorithm.

Definition 1. *Let* $\mathbf{X}\triangleq\{\boldsymbol{\mu},\mathbf{P}\}$ *and* $\mathbf{Y}\triangleq\{\mathbf{m},\mathbf{I}\}$. *We denote the total electricity cost function* $f(\mathbf{X},\mathbf{Y})$, *the workload constraint function* $g(\mathbf{X},\mathbf{Y})$, *and the charging/discharging constraint functions* $h1_i(\mathbf{X},\mathbf{Y})$, $h2_i(\mathbf{X},\mathbf{Y})$, $k1_i(\mathbf{X},\mathbf{Y})$, *and* $k2_i(\mathbf{X},\mathbf{Y})$ *as follows:*

$$f(\mathbf{X},\mathbf{Y}) = \sum_{i=1}^{N}\{(S_i+VC_i)P_i+VC_bI_i+VC_iH_i\},$$

$$g(\mathbf{X},\mathbf{Y}) = \sum_{j=1}^{K}A_j+\sum_{i=1}^{N}\frac{1}{W_{i,max}}-\sum_{i=1}^{N}m_i\mu_i,$$

$$h1_i(\mathbf{X},\mathbf{Y}) = P_i - P_{i,max}I_i, \forall i\in\{1,2,\ldots,N\},$$

$$h2_i(\mathbf{X},\mathbf{Y}) = -P_{i,min}I_i - P_i, \forall i\in\{1,2,\ldots,N\},$$

$$k1_i(\mathbf{X},\mathbf{Y}) = -P_i - H_i, \forall i\in\{1,2,\ldots,N\},$$

$$k2_i(\mathbf{X},\mathbf{Y}) = P_i + H_i - G_{i,max}, \forall i\in\{1,2,\ldots,N\}.$$

Definition 2. *Let*

$$\Theta = \{m_i\in[0,M_i], I_i\in\{0,1\} \mid \exists\mu_i\in[0,\mu_{i,max}], P_i \text{ such that } g(\mathbf{X},\mathbf{Y})\leq 0, h1_i(\mathbf{X},\mathbf{Y})\leq 0, h2_i(\mathbf{X},\mathbf{Y}) \leq 0, k1_i(\mathbf{X},\mathbf{Y})\leq 0, k2_i(\mathbf{X},\mathbf{Y})\leq 0, \forall i\}.$$

For any fixed $\hat{\mathbf{Y}} = \{\hat{\mathbf{m}}, \hat{\mathbf{I}}\} \in \Theta$, *we define the subproblem NLP(*$\hat{\mathbf{Y}}$*) as follows:*

$$\min_{\mathbf{X}} f(\mathbf{X}, \hat{\mathbf{Y}})$$

s.t.

$$\begin{aligned} &g(\mathbf{X}, \hat{\mathbf{Y}}) \leq 0, \\ &h1_i(\mathbf{X}, \hat{\mathbf{Y}}) \leq 0, \forall i \in \{1, 2, \ldots, N\}, \\ &h2_i(\mathbf{X}, \hat{\mathbf{Y}}) \leq 0, \forall i \in \{1, 2, \ldots, N\}, \\ &k1_i(\mathbf{X}, \hat{\mathbf{Y}}) \leq 0, \forall i \in \{1, 2, \ldots, N\}, \\ &k2_i(\mathbf{X}, \hat{\mathbf{Y}}) \leq 0, \forall i \in \{1, 2, \ldots, N\}. \end{aligned}$$

Definition 3. *Let*

$$\Lambda = \left\{ \boldsymbol{\psi} \in \mathcal{R}^{4N+1} : \boldsymbol{\psi} \succeq 0 \text{ and } \sum_{i=1}^{4N+1} \psi_i = 1 \right\}.$$

Define

$$\begin{aligned} J(\mathbf{X}, \mathbf{Y}, \boldsymbol{\psi}) &\triangleq \psi_1 g(\mathbf{X}, \mathbf{Y}) \\ &+ \sum_{i=1}^{N} \{\psi_{i+1} h1_i(\mathbf{X}, \mathbf{Y}) + \psi_{i+N+1} h2_i(\mathbf{X}, \mathbf{Y}) \\ &+ \psi_{i+2N+1} k1_i(\mathbf{X}, \mathbf{Y}) + \psi_{i+3N+1} k2_i(\mathbf{X}, \mathbf{Y})\}. \end{aligned}$$

For any $\hat{\mathbf{Y}} \in \Theta$, *we define the feasibility-check problem NLPF(*$\hat{\mathbf{Y}}$*) as follows:*

$$\min_{\mathbf{X}} \theta$$

s.t.

$$\theta \geq J(\mathbf{X}, \hat{\mathbf{Y}}, \boldsymbol{\psi}), \quad \forall \boldsymbol{\psi} \in \Lambda.$$

Definition 4. *Let* $\boldsymbol{\kappa} \in \mathcal{R}^{4N+1}$ *and* $\boldsymbol{\kappa} \succeq 0$. *Define*

$$\begin{aligned} L(\mathbf{X}, \mathbf{Y}, \boldsymbol{\kappa}) &\triangleq f(\mathbf{X}, \mathbf{Y}) + \kappa_1 g(\mathbf{X}, \mathbf{Y}) \\ &+ \sum_{i=1}^{N} \{\kappa_{i+1} h1_i(\mathbf{X}, \mathbf{Y}) + \kappa_{i+N+1} h2_i(\mathbf{X}, \mathbf{Y}) \\ &+ \kappa_{i+2N+1} k1_i(\mathbf{X}, \mathbf{Y}) + \kappa_{i+3N+1} k2_i(\mathbf{X}, \mathbf{Y})\}. \end{aligned}$$

The master problem is stated as follows:

$$\min_{\mathbf{Y}, z_0} z_0$$

s.t.

$$\begin{aligned} z_0 &\geq \min_{\mathbf{X}} L(\mathbf{X}, \mathbf{Y}, \boldsymbol{\kappa}), \quad \forall \boldsymbol{\kappa} \succeq 0, \\ 0 &\geq \min_{\mathbf{X}} J(\mathbf{X}, \mathbf{Y}, \boldsymbol{\psi}), \quad \forall \boldsymbol{\psi} \in \Lambda. \end{aligned}$$

Definition 5. *The relaxed master problem RMGBD* (p, q) *is stated as follows:*

$$\min_{\mathbf{Y}, z_0} z_0$$

s.t.

$$\begin{aligned} z_0 &\geq \min_{\mathbf{X}} L(\mathbf{X}, \mathbf{Y}, \boldsymbol{\kappa}^i), \quad \forall i \in \{1, 2, \ldots, p\}, \\ 0 &\geq \min_{\mathbf{X}} J(\mathbf{X}, \mathbf{Y}, \boldsymbol{\psi}^j), \quad \forall j \in \{1, 2, \ldots, q\}. \end{aligned}$$

Based on the definitions above, an iterative algorithm based on the generalized Benders decomposition technique is described as follows.

Algorithm 2: Generalized Benders Decomposition to **P4**

Input: An initial feasible solution $\mathbf{Y}^1$ and a convergence tolerant parameter $\epsilon > 0$
Output: An optimal solution $(\mathbf{X}^*, \mathbf{Y}^*)$ to **P4**

```
LB^0 = -∞, UB^0 = +∞ ;
p = q = 0, k = 1 ;
foreach iteration k do
    Solve the NLP(Y^k);
    if NLP(Y^k) is feasible then
        Obtain an optimal solution X^k and an optimal multiplier vector κ*;
        p ← p + 1;
        κ^p = κ*;
        UB^k = min{UB^{k-1}, f(X^k, Y^k)};
        if UB^k = f(X^k, Y^k) then
            (X*, Y*) = (X^k, Y^k);
    else
        Solve the NLPF(Y^k);
        Obtain an optimal solution X^k and an optimal multiplier vector ψ*;
        q ← q + 1;
        ψ^q ← ψ*
    Solve the RMGBD (p, q);
    Obtain an optimal solution (Y^{k+1}, z_0^*);
    LB^k = z_0^*;
    if LB^k + ϵ ≥ UB^k then
        Stop and return (X*, Y*);
    else
        k ← k + 1;
```

5 Algorithmic Performance Analysis

In this section, we analyze the feasibility and algorithmic performance of our algorithm. To start with, we define an upper bound V_{max} on parameter V as follows:

$$V_{max} = \min_i \frac{E_{i,max} - P_{i,max} - P_{i,min}}{C_{i,max} - C_{i,min}}. \tag{26}$$

Next, the optimal solution to **P3** has the following properties that are useful for the following analysis of algorithmic performance.

Lemma 2. *The optimal solution to P3 has the following properties:*

1. *If* $S_i(t) > -VC_{i,min}$, *the optimal solution always choose* $P_i^*(t) \leq 0$.
2. *If* $S_i(t) < -VC_{i,max}$, *the optimal solution always choose* $P_i^*(t) \geq 0$.

Proof. For each data center D_i and time period t,

1. When $S_i(t) > -VC_{i,min}$, suppose $P_i^*(t) > 0$, then we have $G_i^*(t) > m_i^*(t)H_i^*(t)$. According to the objective of **P3**, in this case, the value of the objective should always be larger than the case that $P_i(t) = 0$ and $G_i(t) = m_i^*(t)H_i^*(t)$ where $m_i^*(t)$ and $H_i^*(t)$ do not change. This results in the contradiction because our algorithm is always trying to minimize the objective function. Hence, when $S_i(t) > -VC_{min}$, $P_i(t)$ cannot be strictly greater than zero, i.e., the battery would not charge.
2. When $S_i(t) < -VC_{i,max}$, suppose $P_i^*(t) < 0$, then we have $G_i^*(t) < m_i^*(t)H_i^*(t)$. Similarly, according to the objective of **P3**, in this case, the value of the objective should be always larger than the case that $P_i(t) = 0$ and $G_i(t) = m_i^*(t)H_i^*(t)$ where $m_i^*(t)$ and $H_i^*(t)$ do not change. This results in the contradiction because our algorithm is always trying to minimize the objective function. Hence, when $S_i(t) < -VC_{i,max}$, $P_i(t)$ cannot be strictly less than zero, i.e., the battery would not discharge. □

Then, we have the following theorem about the algorithmic performance of our proposed algorithm:

Theorem 2. *Suppose the initial battery energy level $E_{i,ini} \in [0, E_{i,max}]$. Implementing the above algorithm with any fixed parameter $V \in (0, V_{max}]$ for all periods, we have the following performance guarantees for each data center D_i:*

1. *The battery energy level $E_i(t)$ is always in the range $[0, E_{i,max}]$ for all time slots t.*
2. *All control decisions are feasible.*
3. *If $\mathbf{A}(t)$ and $\mathbf{C}(t)$ are i.i.d. over slots, then the time-average expected electricity cost under our algorithm is within bound B/V of the optimal value:*

$$\limsup_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N} \mathbb{E}\{G_i(t)C_i(t) + \mathbf{1}_{\{P_i(t)\neq 0\}}C_b\} \\ \leq Q^{OPT} + B/V, \tag{27}$$

where B is a constant given by

$$B \triangleq \sum_{i=1}^{N} \frac{\max\{P_{i,min}^2, P_{i,max}^2\}}{2}. \tag{28}$$

In the following, we prove **Theorem 2**.

Proof.

1. To show $0 \leq E_i(t) \leq E_{i,max}$, according to the definition of $S_i(t)$, it is equivalent to show that for each data center D_i,

$$S_i(t) \geq -VC_{i,max} - P_{i,min}, \tag{29}$$

and

$$S_i(t) \leq E_{i,max} - VC_{i,max} - P_{i,min}. \tag{30}$$

As $0 \leq E_{i,ini} \leq E_{i,max}$, the above inequalities hold for $t = 0$. We prove in the following that this constraint is satisfied for all periods by induction.

Suppose inequalities (29), (30) hold for time period t, we need to show that it also holds for time period $t+1$

a. We first prove $S_i(t+1) \leq E_{i,max} - VC_{i,max} - P_{i,min}$: if

$$-VC_{i,min} < S_i(t) \leq E_{i,max} - VC_{i,max} \\ - P_{i,min},$$

then from Lemma 2, we must have $P_i^*(t) \leq 0$. Using (18) and $P_i(t) \geq -P_{i,min}$, we have $S_i(t+1) \leq S_i(t) \leq E_{i,max} - VC_{i,max} - P_{i,min}$; if $-VC_{i,max} - P_{i,min} \leq S_i(t) \leq -VC_{i,min}$, then from (18) and $P_i(t) \leq P_{i,max}$, we have $S_i(t+1) \leq P_{i,max} - VC_{i,min}$. For any $0 < V \leq V_{max}$, from the definition (26) of V_{max}, we have

$$E_{i,max} - VC_{i,max} - P_{i,min} \geq -VC_{i,min} \\ + P_{i,max} \geq S_i(t+1).$$

Thus, we have $S_i(t+1) \leq E_{i,max} - VC_{i,max} - P_{i,min}$.

b. Then, we prove $S_i(t+1) \geq -VC_{i,max} - P_{i,min}$: if $-VC_{i,max} - P_{i,min} \leq S_i(t) < -VC_{i,max}$, then from Lemma 1, we must have $P_i^*(t) \geq 0$. Using the (18), we have $S_i(t+1) \geq S_i(t) \geq -VC_{i,max} - P_{i,min}$; if

$$-VC_{i,max} \leq S_i(t) \leq E_{i,max} - VC_{i,max} \\ - P_{i,min},$$

then from (18) and $P_i(t) \geq -P_{i,min}$, $S_i(t+1) \geq -VC_{i,max} - P_{i,min}$. From the above discussion, we obtain $S_i(t+1) \geq -VC_{i,max} - P_{i,min}$.

2. As the constraint on $E_i(t)$ for each data center D_i is satisfied as shown in 1 and we make our decisions to satisfy all constraints in **P3**, combining them together, all constraints of **P1** are satisfied. Therefore, our control decisions are feasible to **P1**.
3. As we mentioned before, our algorithm is always trying to greedily minimize the R.H.S. of the upper bound (21) of the drift-plus-penalty term at each period t over all possible feasible control policies including the optimal, stationary policy given in Theorem 1. Therefore, by plugging this policy into the R.H.S. of the inequality (21), we obtain the following:

$$\triangle_V(t) \leq \\ B + V\mathbb{E}\left\{\sum_{i=1}^{N}[G_i^{stat}(t)C_i(t) + \mathbf{1}_{\{P_i(t)\neq 0\}}C_b] \mid \mathbf{S}(t)\right\} \\ = B + VQ^{REL} \leq B + VQ^{OPT}.$$

Taking the expectation of both sides, using the law of iterative expectation, and summing over $t \in \{0, 1, 2, \ldots, T-1\}$, we have

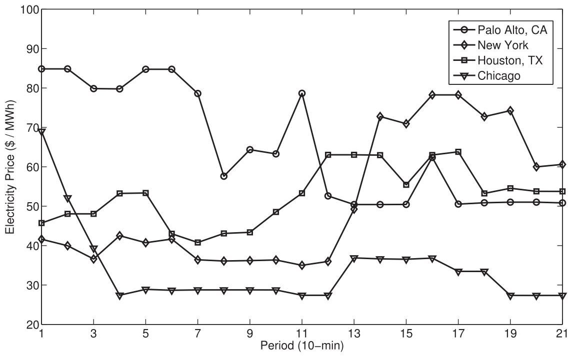


Fig. 2. 10-min average real-time electricity prices in 3 hours at four locations [18], [22].

$$V\sum_{i=1}^{N}\sum_{t=0}^{T-1}\mathbb{E}\{G_i(t)C_i(t)+\mathbf{1}_{\{P_i(t)\neq 0\}}C_b\}$$
$$\leq BT+VTQ^{OPT}-\mathbb{E}\{L(T)\}+\mathbb{E}\{L(0)\}.$$

Diving both side by T, let $T\to\infty$, and using the facts that $E\{L(0)\}$ are finite and $E\{L(t)\}$ are nonnegative, we arrive at the following performance guarantee:

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\mathbb{E}\{G_i(t)C_i(t)+\mathbf{1}_{\{P_i(t)\neq 0\}}C_b\}$$
$$\leq Q^{OPT}+B/V,$$

where Q^{OPT} is the optimal objective value, B is a constant, and V is a control parameter which has the maximum value given by (26). □

6 Numerical Evaluation

In this section, we evaluate the performance of our proposed algorithm based on real-world workload and electricity price data sets. To reduce the overhead between switching the servers on/off across different time periods, the scheduling horizon is divided into discrete time periods with 10 min at each period. We consider a request-response type of cloud service. To accommodate other kinds of typical cloud services such as batch computing and session-based application, only minor modifications need to be made for the traffic distribution constrains and application QoS requirements in our model. We first describe the real-world data sets and system parameters used in this paper. Then, we illustrate the improved energy cost saving of our scheme in comparison with some benchmark schemes. The performance changes due to different battery capacities, battery cost, and QoS requirements are also analyzed.

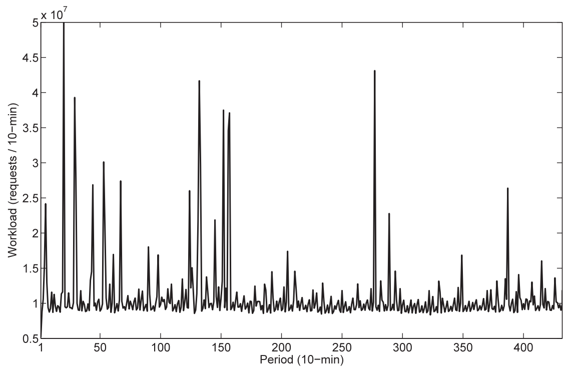


Fig. 3. 10-min average workload in three days [23].

6.1 Experimental Setup

For the purpose of illustration, we consider several simulated data centers where the multiple front-end proxy servers are merged into one. The front-end proxy server acts as a load balancer which receives the incoming traffic requests and distributes workload to different data centers in different locations. It is also responsible for sending control decisions to back-end servers to configure the servers and manage the battery charging/discharging operation. Four different data centers, each having a battery, at different geographic locations are assumed in this evaluation.

Electricity price data. We use the 5-min locational marginal prices (LMP) in real-time electricity markets at four different locations: Chicago, New York, Palo Alto, CA and Houston, TX that host Google's data centers. The data set is obtained from the publicly available government sources [18], [22]. Based on this raw data, we calculate the average electricity price over disjoint 10 minute intervals. The time horizon we consider in this paper is from January 1 to January 21, 2011. In total, this duration includes three weeks or 3,024 10-min periods. A portion of the average 10-min real-time electricity prices during the first week of January 2011 at the four different locations is plotted in Fig. 2. The electricity price is in unit of $/MWh.

Workload data. The real-world workload data we use in the evaluations are a set of I/O traces taken from six RAID volumes at MSR Cambridge [23]. The original traced period is only one week and we repeat it to get a three-week workload traces. Fig. 3 shows the request number variations in different 10-min periods for three days. The peak-to-average ratio of the workload is 4.5.

System parameters. We assume that the servers at one location are homogeneous. Note that our model is quite general and can be easily extended into the heterogenous case with only additional notations. The server parameters at each locations are presented in Table 1, where the service rate (in unit of requests per second) is estimated by the

TABLE 1
Server Parameters in Four Locations

Location	Server	$\mu_{i,max}$ (requests/s)	α_i	β_i (Watt)	γ_i	M_i (servers)
Chicago	AMD Athlon	2	12.5	150	3	15000
New York	AMD Athlon	2	12.5	150	3	10000
Palo Alto, CA	INTEL Pentium 630	1.5	44.44	100	3	10000
Houston, TX	INTEL Pentium 950	2.5	9.6	100	3	10000

average size of the workload request (in unit of bytes per request) as well as CPU and server architecture. We choose $\mathbf{PUE}_i = 1.3$ in all our evaluations to get a conservative estimate of the cost savings. The delay constraint at each data center is chosen to be 1 ms, i.e., $W_{i,max} = 0.001$ s.

6.2 Performance Evaluation

In order to analyze the performance improvement due to our scheme, we compare it with the following schemes that either represent the current practice in data center power management or are the state-of-the-art techniques proposed by previous work [3], [5], [20]: 1) Static load balancing (SLB): current data centers usually runs a constant number of servers to serve the workload. In order to satisfy the time-varying demand, data centers usually overly provision and keep more running servers than what is needed to meet the peak load. In the evaluation, we assume that the data center has complete workload information ahead of time and provisions exactly to satisfy the peak load. Moreover, the amount of workload routed to different data centers is proportional to the service capacity of data centers regardless of electricity price. We assume that all servers are activated at all times. However, the service rates of the servers can be adjusted in every period; 2) Price-aware load balancing (PLB): the scheme is similar to the heuristic proposed by Qureshi et al. [3] that routes more jobs to data centers with lower electricity price. In the evaluation, we assume that the workload is first routed to the data center with the lowest electricity price. Then, we route the remaining workload to the data center with the second lowest price, and so on. Again, all servers are assumed to be activated at all times and service rates can be tuned in every period; 3) Price-aware dynamic provisioning (PDP): this scheme is proposed by Rao et al. [5], [20], which consider both traffic routing and dynamic server provisioning to exploit the spatial variation of electricity price in real-time electricity markets. However, energy storage facilities in data centers are not considered in these work. Therefore, the temporal variation of electricity prices is not utilized.

In the first evaluation, we compare our algorithm with the three benchmark schemes above using the real-world traces. Note that the performance of our scheme depends on the battery capacity and the battery cost. We choose $C_b = 0.1$ \$ and $E_{i,max} = 300$ kWh. The maximum charge and discharge rate are set to be $P_{i,max} = P_{i,min} = 10$ kWh. The initial battery energy level at each data center is chosen to be zero, i.e., $E_{i,ini} = 0$. Let $V = V_{max}$. The result is shown in Fig. 4a. From the figure, we can see that both our scheme and the PDP can get much better performance than the SLB and the PLB because of turning off unnecessary servers rather than leaving them idle. Also, our scheme performs better than the PDP because of the introduction of energy storage, which can charge when the price is low while discharging when the price is high. In the following, we consider the impact of individual parameters on the performance of our scheme compared to the PDP.

Impact of battery capacity. In this evaluation, we vary the battery capacities of data centers with other parameters fixed. We set $E_{i,max} = \{300, 400, 500\}$ kWh and $V = V_{max}$. The result is illustrated in Fig. 4b. From the figure, it is clear that the larger the battery is, the more cost saving our

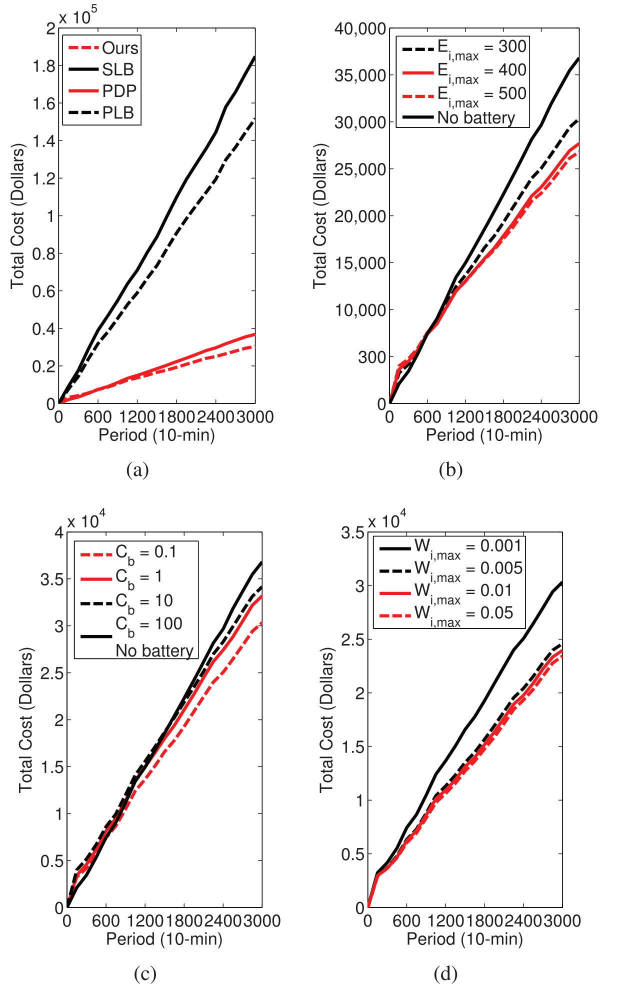


Fig. 4. (a) Comparison of the total energy cost in four approaches. (b) The impact of battery capacity on the cost saving. (c) The impact of battery cost on the cost saving. (d) The impact of QoS requirements on the cost saving.

proposed algorithm can obtain, which coincides with the algorithmic performance results of our algorithm in Theorem 2. As we have mentioned before, the saving comes from the fact that our algorithm would charge the battery when the electricity price is low while discharging it when the electricity price is low.

Impact of battery cost. Currently, the battery is still expansive. The charging or discharging operation would reduce the lifetime of the battery. However, it is expected that the cost of battery would decrease greatly in the next decade. In this evaluation, we estimate the impact of battery cost on the cost saving of our algorithm. We set $C_b = \{0.1, 1, 10, 100\}$ \$ and keep $E_{i,max} = 300$ kWh fixed. The result is shown in Fig. 4c. Note that when the battery cost per operation is very large (e.g., 100 \$), our algorithm would not charge or discharge the battery at all, so it is the same as the scheme in [20]. As the battery cost increases, the total cost saving compared with PDP would decrease since the opportunity to utilize the temporal variation of electricity prices is smaller.

Impact of QoS requirement. In this setting, we adjust the QoS requirements of customer requests while fixing other parameters to see the impact of QoS requirement on the performance of our scheme. We choose $W_{i,max} = \{0.001, 0.005, 0.01, 0.05\}$ s. As observed in Fig. 4d, the increase of the maximum average response time gives more opportunity to optimize the energy cost, since fewer number of servers needs to be turned on to serve the same amount of workload.

7 Conclusion

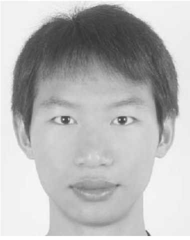
In this paper, we apply the Lyapunov optimization technique to solve the problem of optimal traffic distribution, server configuration, and battery management in data centers for location-varying and time-varying electricity prices under wholesale electricity markets. The algorithm we propose matches the intuition of distributing more traffic into data centers with lower electricity price and charging when electricity price is low while discharging when electricity price is high. Moreover, it is easy to implement online and can give analytic bound on the performance. With the increase of battery capacity, our algorithm can get arbitrarily close to the optimal value. Numerical evaluations based on real-world traces show that our algorithm can result in significant energy cost reduction without scarifying the customer QoS requirements.

Acknowledgments

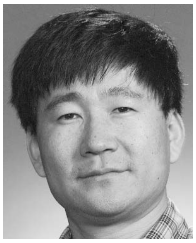
This work was partially supported by the US National Science Foundation (NSF) under grants CNS-0916391, CNS-1147813, and ECCS-1129062. A preliminary version has been published in IEEE GLOBECOM 2011 [1].

References

[1] Y. Guo, Z. Ding, Y. Fang, and D. Wu, "Cutting Down Electricity Cost in Internet Data Centers by Using Energy Storage," *Proc. IEEE GLOBECOM '11,* Dec. 2011.
[2] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A View of Cloud Computing," *Comm. ACM,* vol. 53, no. 4, pp. 50-58, Apr. 2010.
[3] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs, "Cutting the Electric Bill for Internet-Scale Systems," *ACM SIGCOMM Computer Comm. Rev.,* vol. 39, no. 4, pp. 123-134, Aug. 2009.
[4] U. Hoelzle and L.A. Barroso, *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines,* first ed. Morgan & Claypool, 2009.
[5] L. Rao, X. Liu, L. Xie, and W. Liu, "Minimizing Electricity Cost: Optimization of Distributed Internet Data Centers in a Multi-Electricity-Market Environment," *Proc. IEEE INFOCOM '10,* pp. 1-9, Mar. 2010.
[6] M.J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems.* Morgan & Claypool, 2010.
[7] M. Lin, "Algorithmic Issues in Green Data Centers," master's thesis, California Inst. of Technology, 2011.
[8] S. Albers, "Energy-Efficient Algorithms," *Comm. ACM,* vol. 53, no. 5, pp. 86-96, May 2010.
[9] F. Yao, A. Demers, and S. Shenker, "A Scheduling Model for Reduced CPU Energy," *Proc. 36th Ann. Symp. Foundations of Computer Science,* pp. 374-382, Oct. 1995.
[10] K. Pruhs, P. Uthaisombut, and G. Woeginger, "Getting the Best Response for Your Erg," *ACM Trans. Algorithms,* vol. 4, no. 3, pp. 38:1-38:17, July 2008.
[11] D.P. Bunde, "Power-Aware Scheduling for Makespan and Flow," *Proc. 18th Ann. ACM Symp. Parallelism in Algorithms and Architectures (SPAA '06),* pp. 190-196, July 2006.
[12] N. Bansal, H.-L. Chan, and K. Pruhs, "Speed Scaling with an Arbitrary Power Function," *Proc. 20th Ann. ACM-SIAM Symp. Discrete Algorithms (SODA '09),* pp. 693-701, Jan. 2009.
[13] M. Lin, A. Wierman, L.L.H. Andrew, and E. Thereska, "Dynamic Right-Sizing for Power-Proportional Data Centers," *Proc. IEEE INFOCOM '11,* pp. 1098-1106, Apr. 2011.
[14] R. Urgaonkar, B. Urgaonkary, M.J. Neely, and A. Sivasubramaniam, "Optimal Power Cost Management Using Stored Energy in Data Centers," *Proc. ACM Int'l Conf. Measurement and Modeling of Computer Systems (SIGMETRICS '11),* pp. 221-232, June 2011.
[15] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, "Benefits and Limitations of Tapping into Stored Energy for Datacenters," *Proc. 38th Ann. Int'l Symp. Computer Architecture,* pp. 341-352, June 2011.
[16] D.P. Bertsekas and R.G. Gallager, *Data Networks,* second ed. Prentice-Hall, 1992.
[17] A. Gandhi, M. Harchol-Balter, R. Das, and C. Lefurgy, "Optimal Power Allocation in Server Farms," *Proc. 11th Int'l Joint Conf. Measurement and Modeling of Computer Systems (SIGMETRICS '09),* pp. 157-168, Aug. 2009.
[18] Fed. Energy Regulatory Commission, http://www.ferc.gov/, 2012.
[19] D.P. Bertsekas, *Dynamic Programming and Optimal Control,* second ed. Athena Scientific, 2000.
[20] L. Rao, X. Liu, M.D. Ilic, and J. Liu, "Distributed Coordination of Internet Data Centers under Multiregional Electricity Markets," *Proc. IEEE,* vol. 100, no. 1, pp. 269-282, Jan. 2012.
[21] A.M. Geoffrion, "Generalized Benders Decomposition," *J. Optimization Theory and Applications,* vol. 10, no. 4, pp. 237-260, 1972.
[22] United States Energy Information Administration, http://www.eia.gov/, 2012.
[23] D. Narayanan, A. Donnelly, and A. Rowstron, "Write off-Loading: Practical Power Management for Enterprise Storage," *Proc. USENIX Conf. File and Storage Technologies (FAST),* Feb. 2008.



Yuanxiong Guo (S'11) received the BEng degree from the Department of Electronics and Information Engineering, Huazhong University of Science and Technology, Wuhan, China, in 2009. He has been working toward the PhD degree from the Department of Electrical and Computer Engineering at University of Florida, Gainesville since August 2010. His current research interests are in the area of cyber-physical systems including smart grids, sustainable data centers, and cloud computing. He is the recipient of the Best Paper Award from IEEE GLOBECOM 2011, Houston, TX. He is a student member of the IEEE.



Yuguang Fang (F'08) received the PhD degree in systems engineering from Case Western Reserve University in January 1994 and the PhD degree in electrical engineering from Boston University in May 1997. He was an assistant professor in the Department of Electrical and Computer Engineering at New Jersey Institute of Technology from July 1998 to May 2000. He then joined the Department of Electrical and Computer Engineering at University of Florida in May 2000 as an assistant professor, got an early promotion to an associate professor with tenure in August 2003 and to a full professor in August 2005. He holds a University of Florida Research Foundation (UFRF) Professorship from 2006 to 2009, a Changjiang Scholar Chair Professorship with Xidian University, Xi'an, China, from 2008 to 2011, and a Guest Chair Professorship with Tsinghua University, China, from 2009 to 2012. He has published more than 300 papers in refereed professional journals and conferences. He received the US National Science Foundation (NSF) Faculty Early Career Award in 2001 and the Office of Naval Research Young Investigator Award in 2002, and is the recipient of the Best Paper Award in IEEE Globecom (2011), IEEE International Conference on Network Protocols (ICNP, 2006) and the recipient of the IEEE TCGN Best Paper Award in the IEEE High-Speed Networks Symposium, IEEE Globecom (2002). He has also received a 2010-2011 UF Doctoral Dissertation Advisor/Mentoring Award, 2011 Florida Blue Key/UF Homecoming Distinguished Faculty Award and the 2009 UF College of Engineering Faculty Mentoring Award. He is also active in professional activities. He served as the editor-in-chief for *IEEE Wireless Communications* (2009-2012) and serves/served on several editorial boards of technical journals including *IEEE Transactions on Mobile Computing* (2003-2008, 2011-present), *IEEE Network* (2012-present), *IEEE Transactions on Communications* (2000-2011), *IEEE Transactions on Wireless Communications* (2002-2009), *IEEE Journal on Selected Areas in Communications* (1999-2001), *IEEE Wireless Communications Magazine* (2003-2009), and *ACM Wireless Networks* (2001-present). He served on the steering committee for *IEEE Transactions on Mobile Computing* (2008-2010). He has been actively participating in professional conference organizations such as serving as the technical program co-chair for IEEE INOFOCOM 2014, the steering committee co-chair for QShine (2004-2008), the technical program vice-chair for IEEE INFOCOM 2005, the technical program area chair for IEEE INFOCOM (2009-2013), technical program symposium co-chair for IEEE Globecom 2004, and a member of Technical Program Committee for IEEE INFOCOM (1998, 2000, 2003-2008). He is a fellow of the IEEE and a member of the ACM.