Optimal Workload and Energy Storage Management for Cloud Data Centers

Yuanxiong Guo, Yuguang Fang, and Pramod P. Khargonekar
Department of Electrical and Computer Engineering
University of Florida
Gainesville, FL 32611, USA
Email: {guoyuanxiong@, fang@ece., ppk@}ufl.edu

***Abstract*—Electricity expenditure comprises a significant fraction of the total operation cost in cloud data centers and cloud service providers are required to reduce their electricity cost as much as possible. In this paper, we consider electricity cost minimization problem in data centers with energy storage under wholesale electricity markets by considering both the delay-sensitive and delay-tolerant workloads. Considering the stochastic nature of workload arrivals and electricity price, we formulate the problem as a stochastic program and propose an online control algorithm based on the technique of Lyapunov optimization to solve it approximately. The analytical results show that our algorithm can achieve close-to-optimal performance with the increase of energy storage capacities while guaranteeing the quality of service of different workloads.**

***Keywords*-Cloud computing; data center; electricity cost; energy storage; Lyapunov optimization**

I. Introduction

With the growing demand for large-scale computing resources, cloud computing is becoming very popular with different kinds of cloud services ranging from infrastructure-as-a-service, platform-as-a-service, to software-as-a-service [1]. As the key underpinning for supporting these services, more and more data centers are envisioned to be built in the near future. A cloud service provider usually has multiple geographically distributed data centers, each containing hundreds of thousands of servers in order to improve the reliability as well as performance. Power consumption is a critical issue for data centers due to the huge cost of electricity usage as well as the vision to enable the sustainability for information and communication technology.

Motivated by the huge energy consumption, a lot of research has been done toward more energy-efficient data centers during the last decade. Various engineering techniques such as advanced cooling, virtualization, direct current power, multi-core servers, etc., have been employed to improve the power usage effectiveness (PUE) of a data center, which is defined as the ratio of the total power consumption to the IT power consumption [2]. The main focus is on reducing the amount of power usage in data centers. 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 energy amount. Geographical load balancing [3] 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 these 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 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 utility grid 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 maximum power need between 5-30 minutes. This excess energy storage capacity can be used to save electricity cost by the simple intuition of charging when the electricity price is low while discharging when the electricity price is high under wholesale electricity markets. Motivated by [4], [5], our previous work [6] considers minimizing the electricity cost for multiple data centers with energy storage where only delay-sensitive workloads are taken into account. However, in practice, data centers also have other delay-tolerant jobs such as scientific applications, simulations, or MapReduce jobs. This flexibility can be exploited to further reduce the electricity cost by delaying their execution to low price periods. Therefore, in this paper, we consider both the delay-sensitive and delay-tolerant workloads in data centers and propose an online control algorithm based on the Lyapunov optimization techniques [7] to minimize the total electricity cost of a cloud service provider with energy storages under wholesale electricity markets.

The remainder of this paper is organized as follows. Section II describes the models and the problem formulation. An online algorithm is proposed in Section III to solve the problem. Section IV presents the analytical performance

This work was supported in part by the U.S. National Science Foundation under grants CNS-1147813, ECCS-1129061, ECCS-1129062, and Eric Professor endowment at the University of Florida.

978-0-7695-5124-1/13 $31.00 © 2013 IEEE
DOI 10.1109/MILCOM.2013.312

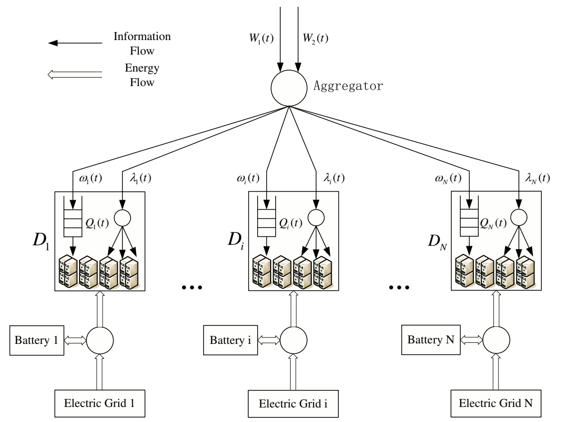

Figure 1. Block diagram for system model

results and Section V concludes the paper.

II. Models and Formulation

We now describe the models we use in this paper. Assume the system is discrete-time with time period matching the timescale at which traffic distribution and charging/discharging decisions can be updated (e.g., 10 min). Further, we assume the traffic inter-arrival time is much shorter than system time period so that traffic distribution can be decided by the average arrival rate during a time period. We consider a cloud service provider having N geographically distributed data centers, denoted by $\mathcal{D} = \{D_1, \ldots, D_N\}$ and a traffic aggregator which is responsible for distributing the total incoming workload to different data centers. 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 Figure 1, which is described in detail as follows.

A. The Workload Model

There are many different workloads in data centers. In general, they can be divided into the following two categories: interactive or transactional applications and non-interactive or batch applications [8]. The interactive workloads such as web services usually process real-time user requests, which have to be completed within a certain time, i.e., there is a maximum response time. Noninteractive batch jobs such as scientific applications, simulations, or MapReduce jobs [9] are often delay-tolerant, which can be scheduled to run at any time as long as the jobs are finished before the deadline, i.e., there is a maximum completion time.

In every period t, interactive workloads and batch workloads arrive at the front-end traffic aggregator. The average arrival rates of interactive and batch workloads are denoted as $W_1(t)$ and $W_2(t)$, respectively. Let $\lambda_i(t)$ denote the average rate of interactive workloads distributed into data center D_i at each period t, $\boldsymbol{\lambda}(t) := (\lambda_1(t), \lambda_2(t), \ldots, \lambda_N(t))$. Then, we have

$$\sum_{i=1}^{N} \lambda_i(t) = W_1(t), \tag{1}$$

$$\lambda_i(t) \geq 0, \forall i. \tag{2}$$

In contrast, the batch workloads can be buffered and serviced later. Denote the average rate of batch workloads distributed into data center D_i as $\omega_i(t)$ at each period t, $\boldsymbol{\omega}(t) := (\omega_1(t), \omega_2(t), \ldots, \omega_N(t))$. Then, we have

$$\sum_{i=1}^{N} \omega_i(t) = W_2(t), \tag{3}$$

$$\omega_i(t) \geq 0, \forall i. \tag{4}$$

B. The Battery Model

We assume that each data center possesses some kind of energy storage, typically UPS batteries. 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 the beginning of 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[1]

$$E_i(t+1) = E_i(t) + P_i(t). \tag{5}$$

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 $P_i(t)$:

$$-P_{i,min} \leq P_i(t) \leq P_{i,max}. \tag{6}$$

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

$$0 \leq E_i(t) \leq E_{i,max}. \tag{7}$$

Note that more complicated model of battery can be easily incorporated into our model without affecting the following analysis, as in [10].

C. The QoS Model

As we have mentioned before, there usually exists a maximum response time indicated in the service level agreement (SLA) between the cloud service provider and its customers. In this paper, we use the $M/GI/1/PS$ queuing model as in [11] to analyze the average response time for interactive workload at data center D_i when the traffic arrival rate is

[1] In this paper, all power quantities such as $P_i(t)$, H_i, $G_i(t)$ are in the unit of energy per period, so the energy produced/consumed in time period t are $P_i(t)$, H_i, $G_i(t)$, respectively.

$\lambda_i(t)$ and there are $m_i(t)$ active servers with service rate μ_i^1 to serve the interactive workloads. Therefore, we have the following delay constraints for the interactive workloads at each data center D_i:

$$\frac{1}{\mu_i^1 - \frac{\lambda_i(t)}{m_i(t)}} \leq d_{i,max}, \tag{8}$$

$$m_i(t) \geq 0, \quad m_i(t) \in \mathbb{N}. \tag{9}$$

For batch workloads at each data center D_i, they are buffered in a queue $Q_i(t)$. Assume that there are $n_i(t)$ active servers used for serving the batch workloads, each with service rate μ_i^2. Then, the dynamics of $Q_i(t)$ is as follows:

$$Q_i(t+1) = \max\{Q_i(t) - n_i(t)\mu_i^2, 0\} + \omega_i(t), \tag{10}$$

$$n_i(t) \geq 0, \quad n_i(t) \in \mathbb{N}. \tag{11}$$

We need to ensure finite average delay for these buffered workloads. Therefore, we have the following QoS requirement for the batch workloads:

$$\overline{\mathbf{Q}} < \infty, \tag{12}$$

where $\overline{\mathbf{Q}}$ is the time average expected queue backlog for the batch workloads and is defined as:

$$\overline{\mathbf{Q}} := \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \mathbb{E}\{Q_i(t)\}. \tag{13}$$

Denote the total number of servers in data center D_i as M_i. We have

$$m_i(t) + n_i(t) \leq M_i, \quad \forall i. \tag{14}$$

D. The Cost Model

In this paper, we are only interested in IT energy consumption without considering cooling consumption. Note that when we reduce the IT energy consumption, the cooling consumption can also be reduced correspondingly. We do not consider power-proportional issue here either, that is, we assume each server at data center D_i consumes either its maximum power H_i when active or zero when inactive. Hence, the energy drawn from the utility grid $G_i(t)$ at data center D_i during time period t is stated as follows:

$$G_i(t) = (m_i(t) + n_i(t))H_i + P_i(t), \quad \forall i. \tag{15}$$

Specifically, when $P_i(t) > 0$, some energy drawn from the grid is used to charge the battery besides powering 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. We assume that the maximum amount of power that can be drawn from the utility grid for data center i in any period is upper bounded by $G_{i,max}$. Thus, we have for all t at each data center i:

$$0 \leq (m_i(t) + n_i(t))H_i + P_i(t) \leq G_{i,max}. \tag{16}$$

As analyzed in [12], the electricity prices in wholesale electricity markets exhibit both spatial and temporal variations. At each data center D_i, we assume a time-varying electricity price $C_i(t)$ in unit of energy with the maximum value $C_{i,max}$ and the minimum value $C_{i,min}$, respectively. Denote $\mathbf{C}(t) = (C_1(t), \ldots, C_N(t))$ as the electricity price vector and $\mathbf{G}(t) = (G_1(t), \ldots, G_N(t))$ 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. Therefore, the total electricity cost of N data centers at time period t is $\sum_{i=1}^{N} C_i(t)G_i(t)$.

E. The Problem Formulation

In this paper, we are interested in minimizing the time-average expected electricity cost. Based on the above models, our problem can be formulated as the following stochastic program, named (P1):

$$\min_{\boldsymbol{\lambda},\boldsymbol{\omega},\mathbf{m},\mathbf{n},\mathbf{P}} F_{av} = \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \mathbb{E}\{G_i(t)C_i(t)\}, \tag{17}$$

subject to constraints (1)-(16).

III. Proposed Algorithm

One challenge to solve the optimization problem above is that the statistics of $W_1(t)$, $W_2(t)$, and $\mathbf{C}(t)$ may be unknown and we need to design an optimal control algorithm under uncertainty. Moreover, the control decisions are coupled due to the time-coupling constraint (7). We adapt the recently developed technique of Lyapunov optimization [7] to design an algorithm to solve it. The algorithm we propose can achieve the range of $O(1/V)$ within the optimal objective value, where V is a parameter constrained by the battery capacity of each data center D_i. A salient feature of our algorithm is that it requires minimum information of the random dynamics in the system and can be easily implemented online.

First, we define a virtual queue $X_i(t)$ as a shifted version of battery energy level $E_i(t)$ in time period t at each data center D_i as follows:

$$X_i(t) = E_i(t) - VC_{i,max} - P_{i,min} \tag{18}$$

where $V > 0$ is a constant control parameter which affects the distance to the optimal value and is constrained by the battery capacity. $\mathbf{X}(t) := (X_1(t), X_2(t), \ldots, X_N(t))$ is used to ensure that the constraint (7) is satisfied in our algorithm as illustrated later. According to (5) of $E_i(t)$, we have the same dynamics for $X_i(t)$ at each data center D_i:

$$X_i(t+1) = X_i(t) + P_i(t). \tag{19}$$

Then, we define a Lyapunov function as follows:

$$L(t) := \frac{1}{2} \sum_{i=1}^{N} [X_i^2(t) + Q_i^2(t)]. \tag{20}$$

Denote $\mathbf{S}(t) := (\mathbf{X}(t), \mathbf{Q}(t))$. Now define the one-period conditional Lyapunov drift as follows:

$$\triangle(t) = \mathbb{E}\{L(t+1) - L(t) \mid \mathbf{S}(t)\}. \quad (21)$$

Here the expectation is taken over the randomness of electricity prices and workload arrivals. Adding a function of the expected electricity cost over one period (i.e., the penalty function) to (21), we obtain the following *drift-plus-penalty* term:

$$\triangle_V(t) := \triangle(t) + V\sum_{i=1}^{N}\mathbb{E}\{G_i(t)C_i(t) \mid \mathbf{S}(t)\}. \quad (22)$$

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*

$$\begin{aligned}\triangle_V(t) \le B &- \sum_{i=1}^{N}\mathbb{E}\{Q_i(t)(n_i(t)\mu_i^2 - \omega_i(t)) \mid \mathbf{S}(t)\} \\ &+ \sum_{i=1}^{N}\mathbb{E}\{X_i(t)P_i(t) \mid \mathbf{S}(t)\} \\ &+ V\sum_{i=1}^{N}C_i(t)\mathbb{E}\{(m_i(t)+n_i(t))H_i + P_i(t) \mid \mathbf{S}(t)\}.\end{aligned} \quad (23)$$

where

$$B := \frac{1}{2}\sum_{i=1}^{N}\left(\max\{P_{i,max}^2, P_{i,min}^2\} + M_i^2(\mu_i^2)^2 + W_{2,max}^2\right). \quad (24)$$

Proof: First, by squaring both sides of (19), we have for each data center D_i:

$$\frac{X_i^2(t+1) - X_i^2(t)}{2} = \frac{P_i^2(t)}{2} + X_i(t)P_i(t). \quad (25)$$

Moreover, we have the following inequality:

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

Similarly, by squaring both sides of (10), and using the fact that $(\max\{Q_i(t) - n_i(t)\mu_i^2, 0\})^2 \le (Q_i(t) - n_i(t)\mu_i^2)^2$, we have for each data center D_i:

$$\begin{aligned}&\frac{Q_i^2(t+1) - Q_i^2(t)}{2} \\ &\le \frac{(n_i(t)\mu_i^2)^2}{2} + \frac{\omega_i^2(t)}{2} - Q_i(t)[n_i(t)\mu_i^2 - \omega_i(t)].\end{aligned} \quad (26)$$

Moreover, we have the following inequality:

$$\frac{(n_i(t)\mu_i)^2}{2} + \frac{\omega_i^2(t)}{2} \le \frac{M_i^2(\mu_i^2)^2 + W_{2,max}^2}{2}.$$

Now by summing (25) and (26) over all data centers, and by defining B as (24), we have:

$$\begin{aligned}&L(t+1) - L(t) \\ &\le B - \sum_{i=1}^{N}Q_i(t)(n_i(t)\mu_i^2 - \omega_i(t)) + \sum_{i=1}^{N}X_i(t)P_i(t).\end{aligned}$$

Taking expectations of both sides above over $\mathbf{S}(t)$ and adding the term $V\sum_{i=1}^{N}\mathbb{E}\{G_i(t)C_i(t) \mid \mathbf{S}(t)\}$ to both sides, we can see that the lemma follows. ■

We now present our algorithm. The main design principle of our algorithm is to choose control actions that approximately minimize the R.H.S. of (23). It is described in Algorithm 1. For each time period t, (P2) is a mixed-integer linear programming (MILP). However, in practice, a data center usually contains thousands of servers, of which a large fraction are active. Hence, we can relax the integer constraint on $m_i(t)$ and $n_i(t), \forall i$, round the resulting solution without significant cost penalties, and get a simple linear optimization problem, which can be solved efficiently in polynomial time using interior-point method [13]. The MILP program can also be solved efficiently by commercial solvers such as CPLEX [14].

Algorithm 1: Workload and Energy Storage Management in Data Centers

foreach *Time period* t **do**

1. Observe the system states $\mathbf{C}(t)$, $W_1(t)$, $W_2(t)$, and $\mathbf{S}(t)$;
2. Choose control decisions $\boldsymbol{\lambda}$, $\boldsymbol{\omega}$, $\mathbf{m}$, $\mathbf{n}$, and $\mathbf{P}$ as the optimal solution to the following optimization problem, called (P2):
   Minimize
   $$\begin{aligned}&\sum_{i=1}^{N}\Big\{(X_i(t) + VC_i(t))P_i(t) + Q_i(t)\omega_i(t) \\ &\quad + (VC_i(t)H_i - Q_i(t)\mu_i^2)n_i(t) + VC_i(t)H_im_i(t)\Big\}\end{aligned}$$
   subject to (1), (2), (3), (4), (6), (8), (9), (11), (14), and (16).
3. Update $X_i(t)$ according to the dynamics (19) and $Q_i(t)$ according to the dynamics (10);

IV. Analytical Performance Analysis

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

$$V_{max} := \min_i \frac{E_{i,max} - P_{i,min} - P_{i,max}}{C_{i,max} - C_{i,min}}. \quad (27)$$

The optimal solution to (P2) has the following property that is useful for the following analysis of algorithmic performance:

***Lemma* 2.** *The optimal solution to (P2) has the following properties:*

- *If $X_i(t) > -VC_{i,min}$, the optimal solution always choose $P_i^*(t) \le 0$.*
- *If $X_i(t) < -VC_{i,max}$, the optimal solution always choose $P_i^*(t) \ge 0$.*

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

1) When $X_i(t) > -VC_{i,min}$, suppose $P_i^*(t) > 0$, then we have $G_i^*(t) > (m_i^*(t) + n_i^*(t))H_i$. According to the objective of (P2), in this case, the value of the objective should always be larger than the case that $P_i(t) = 0$ and $G_i(t) = (m_i^*(t) + n_i^*(t))H_i$ where $m_i^*(t)$ and $n_i^*(t)$ do not change. This results in the contradiction because our algorithm is always trying to minimize the objective function. Hence, when $X_i(t) > -VC_{min}$, $P_i(t)$ cannot be strictly greater than zero, i.e., the battery would not charge.
2) When $X_i(t) < -VC_{i,max}$, suppose $P_i^*(t) < 0$, then we have $G_i^*(t) < (m_i^*(t) + n_i^*(t))H_i$. Similarly, according to the objective of (P2), in this case, the value of the objective should always be larger than the case that $P_i(t) = 0$ and $G_i(t) = (m_i^*(t) + n_i^*(t))H_i$ where $m_i^*(t)$, and $n_i^*(t)$ do not change. This results in the contradiction because our algorithm is always trying to minimize the objective function. Hence, when $X_i(t) < -VC_{i,max}$, $P_i(t)$ 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* 1.** *Assume that $M_i \ge W_{1,max}/(\mu_i^1 - 1/d_{i,max}) + W_{2,max}/\mu_i^2, \forall i$. 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 time periods, we have the following performance guarantees:*

1) *The queue $Q_i(t)$ is deterministically upper bounded by $Q_{i,max}$ for all t and i where*

$$Q_{i,max} := \frac{VC_{i,max}H_i}{\mu_i^2} + W_{2,max}. \tag{28}$$

2) *The battery energy level $E_i(t)$ is always in the range $[0, E_{i,max}]$ for all time periods t.*
3) *If $\mathbf{C}(t)$, $W_1(t)$ and $W_2(t)$ are i.i.d. over periods, then the time-average cost under our algorithm is within bound B/V of the optimal value:*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{N} \mathbb{E}\{G_i(t)C_i(t)\} \le F_{av}^* + B/V \tag{29}$$

where F_{av}^* is the optimal cost obtained by any feasible solution and B is a constant given by (24)*.*

In the following, we prove Theorem 1.

Proof:

1) For each data center D_i, we prove (28) by induction. Clearly, it holds at $t = 0$. Suppose the inequality (28) holds for time period t, we need to show that it also holds for time period $t+1$. If $Q_i(t) \le \frac{VC_{i,max}H_i}{\mu_i^2}$, the maximum increase for $Q_i(t)$ is $W_{2,max}$. Obviously, $Q_i(t+1) \le Q_i(t) + W_{2,max} \le \frac{VC_{i,max}H_i}{\mu_i^2} + W_{2,max}$. If $\frac{VC_{i,max}H_i}{\mu_i^2} < Q_i(t) \le \frac{VC_{i,max}H_i}{\mu_i^2} + W_{2,max}$, according to the objective function in (P2), we know that $VC_i(t)H_i - Q_i(t)\mu_i^2 < 0$, and our algorithm will choose $n_i(t)$ as much as possible. Then, $n_i(t) \ge M_i - W_{1,max}/(\mu_i^1 - 1/d_{i,max}) \ge W_{2,max}/\mu_i^2$. Therefore, the arrival rate of $Q_i(t)$ cannot be greater than the service rate, i.e., the queue length cannot increase. So $Q_i(t+1) \le Q_i(t)$. In summary, $Q_i(t) \le W_{1,max}/(\mu_i^1 - 1/d_{i,max}) + W_{2,max}/\mu_i^2$ for all periods.
2) To show $0 \le E_i(t) \le E_{i,max}$, according to the definition of $X_i(t)$, it is equivalent to show that for each data center D_i,

$$X_i(t) \ge -VC_{i,max} - P_{i,min}, \tag{30}$$

and

$$X_i(t) \le E_{i,max} - VC_{i,max} - P_{i,min}. \tag{31}$$

As $0 \le E_{i,ini} \le E_{i,max}$, the above inequalities holds for $t = 0$. We prove in the following that this constraint is satisfied for all periods by induction. Suppose the inequalities (30) (31) hold for time period t, we need to show that it also holds for time period $t+1$.

- We first prove $X_i(t+1) \le E_{i,max} - VC_{i,max} - D_{i,max}$: if $-VC_{i,min} < X_i(t) \le E_{i,max} - VC_{i,max} - P_{i,min}$, then from Lemma 2, we must have $P_i^*(t) \le 0$. Using (19), we have $X_i(t+1) \le X_i(t) \le E_{i,max} - VC_{i,max} - P_{i,min}$; if $-VC_{i,max} - P_{i,min} \le X_i(t) \le -VC_{i,min}$, then from (19) and $P_i(t) \le P_{i,max}$, we have $X_i(t+1) \le P_{i,max} - VC_{i,min}$. For any $0 \le V \le V_{max}$, from the definition (27) of V_{max}, we have $E_{i,max} - VC_{i,max} - P_{i,min} \ge -VC_{i,min} + P_{i,max} \ge X_i(t+1)$. Thus, we obtain $X_i(t+1) \le E_{i,max} - VC_{i,max} - P_{i,min}$.
- Then we prove $X_i(t+1) \ge -VC_{i,max} - P_{i,min}$: if $-VC_{i,max} - P_{i,min} \le X_i(t) < -VC_{i,max}$, then from Lemma 2, we must have $P_i^*(t) \ge 0$. Using (19), we have $X_i(t+1) \ge X_i(t) \ge -VC_{i,max} - P_{i,min}$; if $-VC_{i,max} \le X_i(t) \le E_{i,max} - VC_{i,max} - P_{i,min}$, then from (19) and $P_i(t) \ge -P_{i,min}$, $X_i(t+1) \ge -VC_{i,max} - P_{i,min}$. From the above discussion, we get $X_i(t+1) \ge -VC_{i,max} - P_{i,min}$.

3) As we have mentioned before, our algorithm is always trying to greedily minimize the R.H.S. of the upper bound (23) of the *drift-plus-penalty* term at each period t over all possible feasible control policies. According to the framework of Lyapunov optimization

[7], there exists a stationary and randomized policy that can satisfy the above constraints while providing the following guarantees:

$$\mathbb{E}\{P_i(t)\} = 0, \forall i \tag{32}$$

$$\mathbb{E}\{n_i(t)\mu_i^2\} \geq \mathbb{E}\{\omega_i(t)\}, \forall i \tag{33}$$

$$\mathbb{E}\{F_{av}(t)\} = \hat{F}_{av}^*. \tag{34}$$

Note that this policy can only be derived based on detailed statistics, which usually has the problem of "curse of dimensionality" if solved by dynamic programming. Moreover, this policy may not be feasible to the original optimization problem because of the constraint (7). Note that we have $\hat{F}_{av}^* \leq F_{av}^*$ due to less constrained for the stationary control policy. In the following, we use the existence of such a policy to derive the performance bound of our proposed algorithm. By substitute this policy into the R.H.S. of (23), we obtain the following:

$$\triangle_V(t) \leq B + V\hat{F}_{av}^* \leq B + VF_{av}^*.$$

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

$$V\sum_{i=1}^{N}\sum_{t=0}^{T-1}\mathbb{E}\{C_i(t)G_i(t)\}$$
$$\leq BT + VTF_{av}^* - \mathbb{E}\{L(T)\} + \mathbb{E}\{L(0)\}.$$

Dividing both sides 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:

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{N}\mathbb{E}\{C_i(t)G_i(t)\} \leq F_{av}^* + B/V,$$

where F_{av}^* is the optimal objective value, B is a constant, and V is a control parameter which has the maximum value given by (27).

■

V. Conclusion

In this paper, we adapt the Lyapunov optimization technique to solve the problem of optimal workload and energy storage management in cloud data centers under wholesale electricity markets. As shown in the analytical performance results, our algorithm can get arbitrarily close to the optimal value with the increase of battery capacities. In the future, we plan to consider renewable-powered data centers as well as doing numerical evaluations based on real-world data traces.

References

[1] 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," *Communications of the ACM*, vol. 53, no. 4, pp. 50–58, Apr. 2010.

[2] U. Hoelzle and L. A. Barroso, *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines*, 1st ed. Morgan & Claypool Publishers, 2009.

[3] Z. Liu, M. Lin, A. Wierman, S. Low, and L. Andrew, "Greening geographical load balancing," in *ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS'11)*, San Jose, CA, Jun. 2011, pp. 233–244.

[4] R. Urgaonkar, U. C. Kozat, K. Igarashi, and M. J. Neely, "Dynamic resource allocation and power management in virtualized data centers," in *Network Operations and Management Symposium (NOMS), 2010 IEEE*, Osaka, Japan, Apr. 2010, pp. 479–486.

[5] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, "Benefits and limitations of tapping into stored energy for datacenters," in *Proceedings of the 38th annual international symposium on Computer architecture*, San Jose, CA, Jun. 2011, pp. 341–352.

[6] Y. Guo, Z. Ding, Y. Fang, and D. Wu, "Cutting down electricity cost in internet data centers by using energy storage," in *Global Telecommunications Conference (GLOBECOM'11), 2011 IEEE*, Houston, TX, Dec. 2011.

[7] M. J. Neely, *Stochastic Network Optimization with Application to Communication and Queueing Systems*. San Rafael, CA: Morgan & Claypool Publishers, 2010.

[8] 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*, London, UK, Jun. 2012.

[9] J. Dean and S. Ghemawat, "Mapreduce: Simplified data processing on large clusters," in *6th Symposium on Operating Systems Design and Implementation (OSDI'04)*, San Francisco, CA, Dec. 2004.

[10] R. Urgaonkar, B. Urgaonkary, M. J. Neely, and A. Sivasubramaniam, "Optimal power cost management using stored energy in data centers," in *ACM International Conference on Measurement and Modeling of Computer Systems (SIGMETRICS'11)*, San Jose, CA, Jun. 2011, pp. 221–232.

[11] M. Lin, A. Wierman, L. L. H. Andrew, and E. Thereska, "Dynamic right-sizing for power-proportional data centers," in *Proc. of IEEE Conference on Computer Communications, INFOCOM 2011*, Shanghai, China, Apr. 2011, pp. 1098–1106.

[12] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs, "Cutting the electric bill for internet-scale systems," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 4, pp. 123–134, Aug. 2009.

[13] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.

[14] IBM ILOG CPLEX Optimizer V12.3. [Online]. Available: http://www-01.ibm.com/software/integration/optimization/cplex-optimizer/