

# MEC-IDC: Joint Load Balancing and Power Control for Distributed Internet Data Centers

Lei Rao  
Department of EI, WNLO  
Huazhong University of S&T  
Wuhan, China  
raolei@smail.hust.edu.cn

Marija Ilic  
Department of ECE  
Carnegie Mellon University  
Pittsburgh, USA  
milic@ece.cmu.edu

Xue Liu  
Department of CSE  
University of Nebraska-Lincoln  
Lincoln, USA  
xueliu@cse.unl.edu

Jie Liu  
Microsoft Research  
Microsoft Corp.  
Redmond, USA  
liuj@microsoft.com

## ABSTRACT

Internet Data Center (IDC) supports the reliable operations of many important Internet on-line services. As the demand on Internet services and cloud computing keep increasing in recent years, the power usage associated with IDC operations has been uprising significantly. The cyber and physical aspects of IDCs interact with each other, and brings unprecedented challenges in power management. While most existing research focuses on reducing power consumptions of IDCs, this paper studies the problem of minimizing the total electricity cost geared to quality of service constraint as well as the location diversity and time diversity of electricity price under multiple electricity markets. We jointly consider both the cyber and physical management capabilities of IDCs, and exploit both the center-level load balancing, and the server-level power control in a unified scheme. We model the problem as a constrained mixed integer programming based on Generalized Benders Decomposition (GBD) technique. Extensive evaluations based on real-life electricity price data for multiple IDC locations demonstrates the effectiveness of our scheme.

## Categories and Subject Descriptors

C.4 [Performance of Systems]: Design studies

## Keywords

Cyber-physical system, distributed Internet data center, load balancing, power control

## 1. INTRODUCTION

Our daily lives are increasingly dependent upon both the

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ICCPs'10 April 13-15, 2010, Stockholm, Sweden.

Copyright 2010 ACM 978-1-4503-0066-7/04/10 ...\$10.00.

world's physical infrastructure and cyber infrastructure. Physical devices, computers, servers, mobile devices are more and more inter connected through networks. These complex systems are referred to as cyber-physical systems (CPSs), as they feature the tight integration and coordination between the system's computational and physical components.

The past decade witnessed tremendous growth of online applications and services, including Email, blogging, social networking, entertainment, and e-commerce. Together with the recent trend of cloud computing, more and more data and computation are migrated to or hosted on the Internet Data Centers (IDCs). IDC has many benefits including data reliability, ease of management for end users, and low amortized cost of ownership. For example, it uses geographical distribution and replication to improve data reliability [14]. Also, IDC operators such as Google, Microsoft, and Amazon.com etc can use economies of scale and statistical multiplexing to amortize the total cost of ownership over a large number of machines and users [1, 10, 18, 20].

Today's IDCs are achieving significant advances in communication and computation capabilities. However, along with the increasing demand of online services and the associated deployment of IDCs, the cost on powering up (and cooling) these Internet Data Centers is increasing significantly. In the U.S., it is reported that IDC facilities consume approximately 2% of all electricity, with an estimated growth rate of 12% per year. A recent study estimated that the worldwide spending on enterprise and data center power and cooling to top \$30 billion in year 2008 and is likely to even surpass the spending on new server hardware [25]. Therefore, the need for power management of IDCs is becoming ever more urgent and important.

On the other hand, the technology advances of IDC design bring many new opportunities for IDC power management. In the server level, existing technics include Chip Multi-Processing (CMP), Dynamic Voltage and Frequency Scaling (DVFS), Sleep Scheduling and Virtual Machine Management. In the data center level, we can exploit load balancing and task scheduling to support power efficiency. These new technologies enable the the complex interactions between the cyber and physical aspects of IDCs as well as bring many new challenges [14]. It is desirable to co-design the software architecture together with the physical infras-

structure to maximize resource utilization and save power. For example, when service load increases, we can leverage load balancing and service aggregation on the computation and communication (cyber) side together with server-level adaptation and sleep (ON/OFF) scheduling on the physical side to save power and at the same time meet service-level agreement (SLA).

There are extensive existing research on power management of server farms [8, 12, 13, 25]. Most of these power management approaches focus on how to reduce the power demands of data center. However, the ultimate goal of IDC service providers (such as Google, Microsoft etc.) is to reduce the total operating cost: it not only depends on the power consumed by the IDCs, but also depends on the electricity price which could exhibit location and time diversity. Recently, there are research work on this goal (i.e. minimizing the total electricity cost) [4, 16]. However, those work did not take server-level design into considerations, which is an important mechanism to reduce power consumption.

In this paper, we propose MEC-IDC, a joint load balancing and power control scheme for distributed IDCs. In our scheme, load balancing is conducted on the data center level, and power control is conducted on the individual server level by exploiting server ON/OFF and Dynamic Voltage and Frequency Scaling (DVFS). The center-level load balancing and server-level power control work together to minimize the total electricity cost while at the same time satisfies the quality of service delay requirements to the users. We formulate the total electricity cost minimization problem for IDCs in a multi-electricity-market environment [16] as a constrained optimization problem, in which the constraints capture the delay guarantee and workload requirements. We then propose an efficient solution based on the extension of Generalized Benders Decomposition (GBD) [2] technique.

The main contribution of this paper is summarized as follows:

- First, we investigate the important problem of minimizing the total electricity cost for IDC service providers under a multi-electricity-market environment. Our formulation captures both center-level load balancing and server-level power control capabilities so as to achieve the goal of minimizing the total electricity cost of IDCs, while in the same time can guarantee the quality of service experienced by end users.
- Second, we propose an efficient solution based on the extension of GBD technique to obtain the optimal load balancing and power control scheme for electricity cost minimization. Further, we evaluate the formulation and solution based on real-life electricity price data and shows that significant total electricity cost reduction can be achieved.

The rest of this paper is organized as follows. Section 2 discusses related work. Section 3 proposes the MEC-IDC design and gives the problem modeling and formulation. Section 4 gives the solution for the two-component design. Section 5 discusses a multi-electricity market environment to emulate the main Google IDC locations and shows that the proposed MEC-IDC scheme could greatly reduce the total electricity cost. Finally, Section 6 concludes the paper.

## 2. RELATED WORK

Cyber-Physical System (CPS) has many applications such as health care, transportation and automotive networks, aerospace and avionics, automated manufacturing, blackout-free electricity generation and distribution, optimization of energy consumption in buildings and vehicles, critical infrastructure monitoring, disaster response, efficient agriculture, and environmental science [17, 34, 7]. Good overviews of different CPS research aspects can be found in the position papers published in the recent NSF workshops on CPSs [22].

Internet Data Center (IDC) is one important emergent CPS. Online applications and services supported by IDC are primarily for communication and computation purposes, but the continuous operation and correct functioning are constrained by the physical resources such as power consumptions. In this paper, we study the total electricity cost minimization problem in a multi-electricity-market environment for IDCs. In this section we discuss the work most relevant to ours. For related work on multi-electricity-market environment, please refer to [16].

### 2.1 Power Management in IDCs

Along with the surging energy usage of IDCs, power management is becoming an ever more important and active research area. A latest overview of challenges toward power management in IDCs is presented in [14]. Energy-efficient servers can be enabled by existing energy saving solutions including low power chipset with chip multiprocessing [15], virtualization technologies [26, 23], power control theories [35] and other techniques [25].

While most of the existing technologies have been developed to reduce the total power consumptions of IDCs [8, 12, 13, 21, 27, 31, 36], recently Qureshi et al. [4] and Rao et al. [16] study the novel problem of minimizing IDC total electricity cost. In [4] Qureshi et al. focus on the evaluation of electricity price data and in [16] Rao et al. focus on modeling and formulation of load balancing for IDCs. However, they do not take server-level design into consideration, which is important for reducing power consumption.

### 2.2 Constrained Mixed Integer Nonlinear Programming: Problem and Solution

The decision variables in constrained mixed integer nonlinear programming problems (MINLP) include both integer variables and continuous variables. Optimization models of an MINLP structure arise in a variety of fields, including chemical engineering and reliability networks. A general MINLP is known to be NP-hard and no efficient solutions exist when the problem size is large because the search space may increase exponentially. However, for some practical applications, the MINLP problem often possesses some special structures that can be exploited for designing effective solutions. One special situation is that fixing the discrete variables renders the objective function convex for continuous variables. In this paper, we utilize the efficient Generalized Benders Decomposition (GBD) [2] technique to solve the mixed integer non-linear programming problem (MINLP) under our formulation. GBD has been used in a variety of research work [6]. The proposed algorithm is expected to converge to the optimal solution within a finite number of iterations.

### 3. TOTAL ELECTRICITY COST MINIMIZATION: DESIGN, MODELING AND FORMULATION

In this section, we first provide an overview of our two-component decoupling design. We then present the details for modeling the total electricity cost, workload constraint and end-to-end delay constraint as well as formulating the total electricity cost minimization problem as a constrained mixed-integer programming problem for Internet Data Centers (IDCs). Based on detailed discussions, we decompose the constrained optimization problem into two subproblems, which coincides with our proposed two-component decoupling design.

#### 3.1 Two-Component Decoupling Design

In this subsection, we provide a description of our two-component decoupling design architecture. Figure 1 shows the overall architecture of total electricity cost minimization problem. There are  $N$  IDC locations and  $C$  front-end web portal servers. Under the dynamics of electricity prices, we calculate the CPU frequency  $f_i$ , the number of On-state servers at IDC location  $i$  as well as the workload  $\lambda_{ji}$  assigned from front-end web portal server  $j$  to IDC location  $i$ .  $L_j$  is the requests demand at each front-end web portal server  $j$ .

The service delay of a request is composed of two parts: the first part is the transmission delay from the front-end web portal to the server in the IDC which services the request; the second part is the delay of service (including the queueing delay) at the server. For most of current deployments, the transmission delay is incurred on the Internet and is usually not directly controllable by the IDC service providers, while the service delay of the servers can be controlled by the service providers. In order to meet the delay quality of service, we design a two-component decoupling approach which takes both transmission delay and server service delay into accounts.

The two-component design is illustrated in Figure 2: after the client request reaches the service provider's front-end web portal, the first component will assign the request to a specific sever in one of the IDC according to the capacities of the IDCs. The request will then be serviced at the assigned server with controllable service speed which is managed by DVFS policies. This corresponds to the second component. The delay from the first to the second component is the transmission delay, while the delay in the second component is the server service delay.

#### 3.2 Notations

We introduce notations used in this paper in Table 1.

#### 3.3 Modeling and Formulation of Total Electricity Cost Minimization Problem

**Formulation of Total Electricity Cost Minimization Problem.** The objective of the power management problem in this paper is to minimize the total electricity cost  $T_{total}$  for IDCs in a multi-electricity-market environment. Hence, we use the total electricity cost <sup>1</sup> as the objective function

<sup>1</sup>Note that energy is the time integral of power consumption, and electricity cost is the time integral of the product of electricity price and power consumption. In this paper we mainly discuss the optimization at a specific time, hence we omit the time integral in the formulation. So we use power

Table 1: Notations

$i$	IDC location $i$
$j$	Front-end web portal server $j$
$t$	Time $t$
$N$	Total number of IDC locations
$C$	Total number of front-end web portal servers
$L_j$	Total workload at front-end web portal server $j$
$Pr_i(t)$	Spot price of electricity at $i$
$Po_i$	Power consumption for one server at $i$
$\lambda_{ji}$	The request rate from $j$ to $i$
$D_i$	The end-to-end delay constraint at $i$
$m_i$	The number of turned on servers at $i$
$f_i$	The CPU frequency at $i$
$A_i$	Server power consumption specified parameter
$B_i$	Server power consumption specified parameter

to be minimized. The total electricity cost function is in the general form:

$$T_{total} = \sum_{i=1}^N m_i Pr_i(t) Po_i.$$

In this paper, we adjust the On/Off of servers and develop a dynamic voltage and frequency scaling (DVFS) scheme to achieve lower power consumption at each IDC location and balance the workloads distributed among IDC locations and front-end web portal servers. Hence, the decision variables are the number of On-state servers  $m_i$ , the server working frequency  $f_i$  and the assigned workloads  $\lambda_{ji}$ . The constraints are the workload constraints from front-end web portal servers and the delay and device constraints at IDC locations.

In conclusion, we have the following optimization problem named **P0**:

$$\min_{m_i, f_i, \lambda_{ji}} \sum_{i=1}^N m_i Pr_i(t) Po_i,$$

subject to

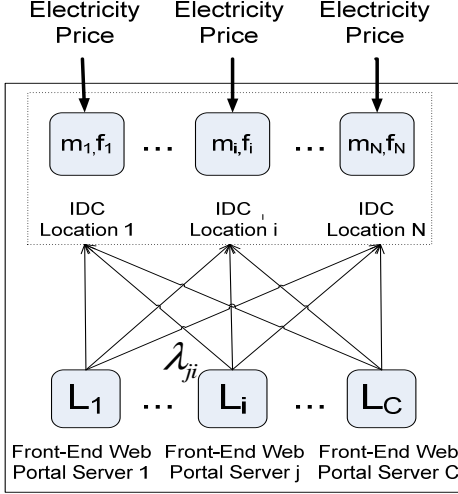
$$\text{Delay constraint, } \quad \forall i = 1, \dots, N,$$

$$\text{Workload constraint, } \quad \forall j = 1, \dots, C,$$

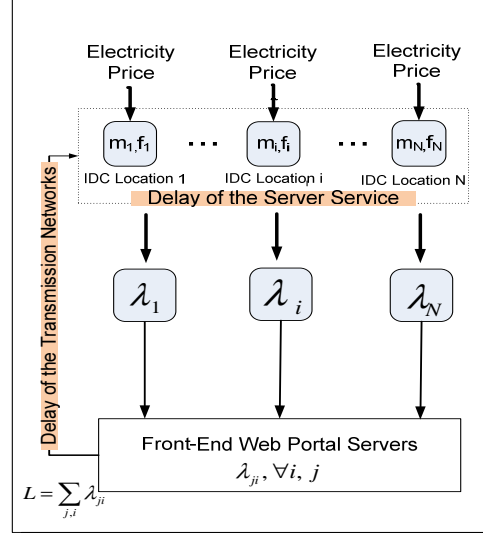
$$\text{Device constraint, } \quad \forall i = 1, \dots, N.$$

**Total Electricity Cost Modeling for IDC.** The electric power grid in North America is operated on a regional basis by regional transmission organizations (RTOs). Regions in which most large-scale data centers are built are under competitive electricity market structure. The spot price of electricity is determined by the clearing of supply and demand functions, while transmission limits are observed. A bottom-up bid-based stochastic price model [28] is adopted here to characterize the hourly spot price dynamics  $Pr_i(t)$ .

In order to simplify the problem, we make a few important assumptions. We assume that the power consumption is equal for all servers at each location. Furthermore, the machines are load-balanced in steady-state, such that each machine at each location has the same utilization, the same frequency, and the same traffic arrival rate. We approximate power consumption  $Po_i$  by the following function of consumption directly instead of energy consumption.



**Figure 1: Overall Architecture of Total Electricity Cost Minimization Problem**



**Figure 2: Decoupling Design Architecture of Total Electricity Cost Minimization Problem**

CPU frequency  $f_i$  for each server at each IDC location  $i$ :

$$Po_i = A_i f_i^p + B_i.$$

where  $A_i$  and  $B_i$  are positive constants. In realistic systems  $p$  varies between 2.5 and 3.  $A_i$ ,  $B_i$ , and  $p$  can be obtained by curve fitting against empirical measurements when profiling the system off-line.

Thus, the total electricity cost for  $N$  IDC locations is given as:

$$T_{total} = \sum_{i=1}^N m_i Pr_i(t)(A_i f_i^p + B_i). \quad (1)$$

**Workload Constraint Modeling.** We denote the requests demand at each front-end web portal server  $j$  as  $L_j$  ( $j = 1, \dots, C$ ). The request arrival rate from front-end web portal server  $j$  to IDC location  $i$  is  $\lambda_{ji}$ . Therefore we have:

$$\sum_{i=1}^N \lambda_{ji} = L_j, \forall j = 1, \dots, C. \quad (2)$$

All the requests demands are sent to several locations of Internet data centers so as to be processed. There are numbers of servers at each location so that it can afford large amount of computations. However, there is a limitation  $M_i$  on the number of servers  $m_i$  at each location  $i$  of IDCs. Therefore, we have

$$m_i \leq M_i, \forall i = 1, \dots, N. \quad (3)$$

**Delay Constraint Modeling.** We use the M/M/n queuing model to model each server in the data center. In the M/M/n queuing model, the average delay  $D$ , is expressed as  $\frac{1}{n\mu - \lambda} P_Q$ , given the number of servers  $n$ , the service rate  $\mu$ , the arrival rate  $\lambda$  and the probability of customers waiting in queue  $P_Q$ . In a data center, without loss of generality, we can assume that the servers are always busy, i.e. there are always requests waiting in queue. Hence, we have  $P_Q$  equals 1.

At IDC location  $i$  with  $m_i$  servers, when each sever has the service rate  $\mu_i$  and the total arrival rate is  $\sum_{j=1}^C \lambda_{ji}$ , the average delay  $D_i^{CPU}$  is given as

$$D_i^{CPU} = \frac{1}{m_i \mu_i - \sum_{i \neq 1}^C \lambda_{ji}}.$$

To meet the requirements of end users, there is a delay constraint  $D_i$  for the average delay at each IDC location  $i$ :

$$0 \leq \frac{1}{m_i \mu_i - \sum_{i \neq 1}^C \lambda_{ji}} \leq D_i.$$

Note that  $\mu_i$  is in terms of  $\#(of\ requests)/s$  and the server working frequency  $f_i$  at each IDC location  $i$  is in terms of  $\#(of\ cycles)/s$ , we convert  $\mu_i$  to  $f_i$  by  $\mu_i = \frac{f_i}{K}$ , where  $K$  is in terms of  $\#(of\ commands)/requests$ . Thus we have:

$$\frac{1}{\frac{m_i f_i}{K} - \sum_{j=1}^C \lambda_{ji}} \leq D_i. \quad (4)$$

Every machine may impose a limit on computing rate due to its physical limitations. We use  $f_{i,max}$  to refer to this limit:

$$f_i \leq f_{i,max}. \quad (5)$$

### 3.4 Decomposition of the Total Electricity Cost minimization Problem

In this section, we present the decomposition of the total electricity cost minimization problem coinciding with the two-component decoupling design shown in subsection 3.1.

Based on the modeling and formulation in the above sub-

section, problem **P0** can be rewritten as follows:

$$\min_{m_i, f_i, \lambda_{ji}} \sum_{i=1}^N m_i Pr_i(t)(A_i f_i^p + B_i),$$

subject to

$$\begin{aligned} \frac{1}{\frac{m_i f_i}{K} - \sum_{j=1}^C \lambda_{ji}} &\leq D_i, \forall i = 1, \dots, N, \\ \sum_{i=1}^N \lambda_{ji} &= L_j, \forall j = 1, \dots, C, \\ m_i &\leq M_i, \forall i = 1, \dots, N, \\ m_i &\in \mathcal{N}, \forall i = 1, \dots, N, \\ f_i &\leq f_{i\_max}, \forall i = 1, \dots, N. \end{aligned} \quad (6)$$

We define  $\lambda_i = \sum_{j=1}^C \lambda_{ji}$ , then we have  $\sum_{i=1}^N \lambda_i = \sum_{j=1}^C L_j = L$ , where  $L$  is the amount of total workload from all front-end web portal servers. Based on the formulation of **P0**, without the loss of generality, we can decompose the load balancing process of planning the workload assigned from each front-end web portal server  $j$  to each IDC location  $i$  into two steps: we first calculate  $\lambda_i$  to decide the amount of workload assigned to each IDC location  $i$  according to  $L$ ; we then determine  $\lambda_{ji}$  based on the relative location information of front-end web portal servers and IDCs. Then we have the equality optimization problem, namely, **P0-1**, as following:

$$\begin{aligned} \sum_{j=1}^C \lambda_{ji} &= \lambda_i, \forall i = 1, \dots, N, \\ \sum_{i=1}^N \lambda_{ji} &= L_j, \forall j = 1, \dots, C. \end{aligned} \quad (7)$$

and the constrained optimization problem, namely **P0-2**, as following:

$$\min_{m_i, f_i, \lambda_i} \sum_{i=1}^N m_i Pr_i(t)(A_i f_i^p + B_i),$$

subject to

$$\begin{aligned} \frac{1}{\frac{m_i f_i}{K} - \lambda_i} &\leq D_i, \forall i = 1, \dots, N, \\ \sum_{i=1}^N \lambda_i &= L, \\ m_i &\leq M_i, \forall i = 1, \dots, N, \\ m_i &\in \mathcal{N}, \forall i = 1, \dots, N, \\ f_i &\leq f_{i\_max}, \forall i = 1, \dots, N. \end{aligned} \quad (8)$$

Let  $x = [f_1, \dots, f_N, m_1, \dots, m_N, \lambda_1, \dots, \lambda_N]^T$  be the vector of decision variables. Introducing the Lagrangian multiplier  $v_2$  and multiplier vectors  $v_1 = [v_{11}, \dots, v_{1N}]^T$ ,  $v_3 = [v_{31}, \dots, v_{3N}]$  and  $v_4 = [v_{41}, \dots, v_{4N}]$ , we can write the La-

grangian function for problem **P0-2** as:

$$\begin{aligned} L(x, v_1, v_2, v_3, v_4) &= \sum_{i=1}^N m_i Pr_i(t)(A_i f_i^p + B_i) \\ &+ \sum_{i=1}^N v_{1i} \left( \frac{1}{D_i} - \frac{m_i f_i}{K} + \lambda_i \right) \\ &+ v_2 \left( L - \sum_{i=1}^N \lambda_i \right) \\ &+ \sum_{i=1}^N v_{3i} (m_i - M_i) \\ &+ \sum_{i=1}^N v_{4i} (f_i - f_{i\_max}). \end{aligned} \quad (9)$$

The Karush-Kuhn-Tucker (KKT) conditions associated with the optimization problem are:

$$\begin{aligned} \nabla_x L(x, v_1, v_2, v_3, v_4) &= 0, \\ \sum_{i=1}^N v_{1i} \left( \frac{1}{D_i} - \frac{m_i f_i}{K} + \lambda_i \right) &= 0, \\ v_2 \left( L - \sum_{i=1}^N \lambda_i \right) &= 0, \\ \sum_{i=1}^N v_{3i} (m_i - M_i) &= 0, \\ \sum_{i=1}^N v_{4i} (f_i - f_{i\_max}) &= 0. \end{aligned} \quad (10)$$

The last four equations are the complementary slackness conditions, which state that for all inactive constraints the corresponding Lagrange multiplier is zero. For our problem, the KKT conditions give:

$$\begin{aligned} \frac{\partial L}{\partial m_i} &= Pr_i(t)(A_i f_i^p + B_i) - v_{1i} \frac{f_i}{K} + v_{3i} = 0, \forall i, \\ \frac{\partial L}{\partial f_i} &= p m_i Pr_i(t) A_i f_i^{(p-1)} - v_{1i} \frac{m_i}{K} = 0, \forall i, \\ \frac{\partial L}{\partial \lambda_i} &= v_{1i} - v_2 = 0, \forall i, j. \end{aligned} \quad (11)$$

Since  $v_{1i} = v_2 \geq 0$ , we have  $\frac{1}{D_i} - \frac{m_i f_i}{K} + \lambda_i = 0, \forall i$ . We can write problem **P0-2** as:

$$\min_{m_i, f_i} \sum_{i=1}^N m_i Pr_i(t)(A_i f_i^p + B_i) T,$$

subject to

$$\begin{aligned} \sum_{i=1}^N \left( \frac{m_i f_i}{K} - \frac{1}{D_i} \right) &= L, \\ m_i &\leq M_i, \forall i = 1, \dots, N, \\ m_i &\in \mathcal{N}, \forall i = 1, \dots, N, \\ f_i &\leq f_{i\_max}, \forall i = 1, \dots, N. \end{aligned} \quad (12)$$

We first solve the above mixed integer nonlinear optimization problem (MINLP) to decide  $m_i$  and  $f_i$  for each IDC location  $i$ , and then calculate the workload assigned to IDC location  $i$  by  $\lambda_i = \frac{m_i f_i}{K} - \frac{1}{D_i}$ .

## 4. SOLUTION AND ALGORITHM DESIGN

In this section, we present the solutions for problem **P0** (including **P0-1** and **P0-2**) in section 3. Based on the decomposition of problem **P0**, we first show that problem **P0-1** can be regarded as an assignment problem. We then solve problem **P0-2**, which is a mixed integer nonlinear programming, by an extension of Generalized Benders Decomposition (GBD) [2] technique.

### 4.1 Solution of Problem P0-1

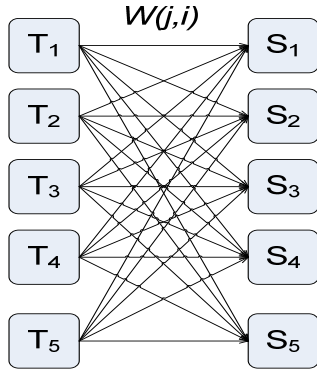
In this subsection, we leverage the relative location information of all the front-end web portal servers and IDCs to solve problem **P0-1**.

We define a general cost function  $W(j, i)$  for front-end web portal server  $j$  and IDC location  $i$ . The function value of  $W(j, i)$  is decided by the delay of transmission networks from front-end web portal server  $j$  to IDC location  $i$ . The problem of choosing  $\lambda_{ji}$  to minimize the total weight function value for all the front-end web portal servers can be described by the following optimization problem:

$$\min_{\lambda_{ji}} \sum_{j=1}^C \sum_{i=1}^N W(j, i) \lambda_{ji},$$

subject to

$$\begin{aligned} \sum_{j=1}^C \lambda_{ji} &= \lambda_i, \forall i = 1, \dots, N, \\ \sum_{i=1}^N \lambda_{ji} &= L_j, \forall j = 1, \dots, C. \end{aligned} \quad (13)$$



**Figure 3: Illustration of the Simple Case in Subsection 4.1**

To illustrate the solution of the above problem, without the loss of generality, we consider a simple case that  $N = 3$  and  $C = 5$ . The above problem can be looked at graphically as shown in Figure 3. We have IDC locations  $S_1, S_2, S_3, S_4$  and  $S_5$  (including the dummy locations) and front-end web portal servers  $T_1, T_2, T_3, T_4$  and  $T_5$ . The lines from  $T_j$  to  $S_i$  represent the values of the cost function, with all the  $W(j, 4)$  and  $W(j, 5)$  set to 0. Given these two sets of nodes with equal size, together with the values of the cost function, problem 13 can be regarded as an assignment problem,

which has been well studied. We can utilize existing algorithms to solve this problem, and interested readers please refer to [11].

### 4.2 Solution and Algorithm Design of Problem P0-2

In this subsection, we extend the efficient Generalized Benders Decomposition (GBD) [2] technique to solve problem **P0-2**, which is a mixed integer non-linear programming problem (MINLP). We first give the basic definitions that will be used in the algorithm:

**DEFINITION 1.** We define the electricity cost function  $F(m, f)$  and the load constraint function  $G(m, f)$  as following:

$$\begin{aligned} F(m, f) &= \sum_{i=1}^N m_i Pr_i(t)(A_i f_i^p + B_i), \\ G(m, f) &= L - \sum_{i=1}^N \left( \frac{m_i f_i}{K} - \frac{1}{D_i} \right). \end{aligned}$$

**DEFINITION 2.** Let  $V = \{m \in M \mid \text{there exists } f \in F \text{ such that } G(m, f) \leq 0\}$ , where  $M = m_i \mid m_i \in (0, M_i], \forall i$  and  $F = f_i \mid f_i \in (0, f_{i,max}], \forall i$ . For any fixed  $\hat{m} \in V$ , we define the subproblem  $NLP(\hat{m})$  as following:

$$\min_{f \in F} F(\hat{m}, f),$$

subject to

$$G(\hat{m}, f) \leq 0.$$

**DEFINITION 3.** For any  $\hat{m} \in F$ , we define the feasibility-check problem  $NLPF(\hat{m})$  as following:

$$\min_{f \in F} \beta,$$

subject to

$$\beta \geq G_t(\hat{m}, f), t = 1, \dots, q.$$

$NLP(\hat{m})$  is feasible if and only if  $NLPF(\hat{m})$  has a positive optimal value  $\beta^* \geq 0$ .

**DEFINITION 4.** Let  $\Lambda = \{\sum_{t=1}^q \psi_t, \psi \geq 0, t = 1, \dots, q\}$ , where  $\psi$  is the optimal multiplier vector for  $NLPF(\hat{m})$  and  $L(f, \kappa) = F(\hat{m}, f) + \kappa^T G(\hat{m}, f)$ , where  $\kappa$  is the optimal multiplier vector for  $NLP(\hat{m})$ . We define the master problem  $MGBD$  as following:

$$\min_{m \in M} \alpha,$$

subject to

$$\begin{aligned} \alpha &\geq \min_{f \in F} L(f, \kappa), \forall \kappa \geq 0, \\ 0 &\geq \min_{f \in F} \psi^T G(\hat{m}, f), \forall \psi \in \Lambda. \end{aligned}$$

Based on the above definitions, an iterative scheme can be developed as follows:

**ALGORITHM 1. Step 0.** Choose  $m^1 \in M$ . Set  $LB^0 = -\infty, UB^0 = +\infty, I^0 = J^0 = \emptyset, k = 1$ .  
**Step 1.** Solve  $NLP(m^k)$ .

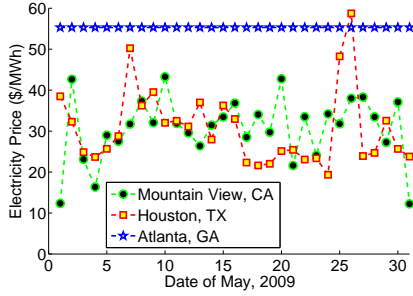


Figure 4: Daily Electricity Price Data for Three Major Locations of Google IDCs on May, 2009

- If  $NLP(m^k)$  is feasible, we obtain an optimal solution  $f^k$  and an optimal multiplier vector  $\kappa^k$ . Set  $I^k = I^{k-1} \cup \{k\}$  and  $J^k = J^{k-1}$ . Set  $UB^k = \min\{UB^{k-1}, F(m^k, f^k)\}$ . If  $UB^k = F(m^k, f^k)$ , set  $(m^*, f^*) = (m^k, f^k)$ .
- If  $NLP(m^k)$  is infeasible, solve  $NLPF(\hat{m})$  and obtain an optimal solution  $f^k$  and an optimal multiplier vector  $\psi^k$ . Set  $J^k = J^{k-1} \cup \{k\}$  and  $I^k = I^{k-1}$ .

**Step 2.** Solve the master problem  $MGBD$  and obtain an optimal solution  $(\alpha^k, m^{\{k+1\}})$ . Set  $LB^k = \alpha^k$ . If  $LB^k \geq UB^k$ , stop and  $(m^*, f^*)$  is the optimal solution to the MINLP. Otherwise, set  $k := k + 1$  and go to **Step 1**.

## 5. PERFORMANCE EVALUATION

In this section, we present the electricity price data at certain locations for Google Internet Data Centers (IDCs) and evaluate the performance of the proposed method based on real-life electricity price data [16]. We show the total electricity cost reductions by comparing results of the proposed total electricity cost minimization method and the optimal power control method based on the data of May 2nd, 2009 [16]. We then present the optimal workload assignments to these IDC locations.

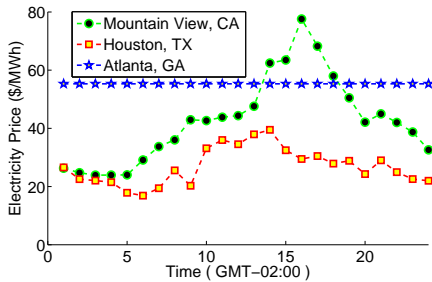


Figure 5: Hourly Electricity Price Data for Three Major Locations of Google IDCs on May 2nd, 2009

### 5.1 Electricity Price at Certain Locations for Google Internet Data Centers

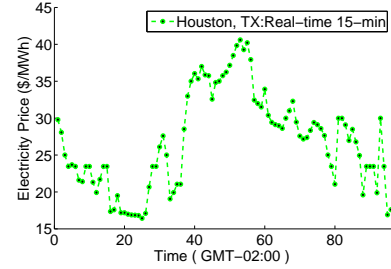
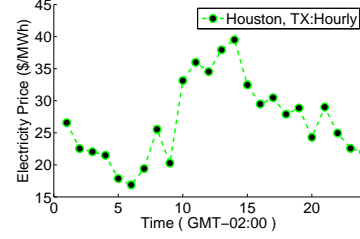


Figure 6: Comparing Price Variation for Hourly Market and Real-time Market in Houston on May 2nd, 2009

In this subsection, we present the electricity price data at certain locations for Google Internet Data Centers (IDCs) obtained from publicly available government agencies [32] [33]. The locations include Mountain View, California, Houston, Texas, and Atlanta, Georgia.

The electricity system in the US is organized in cross-state regions such as New England, PJM (primarily covering Pennsylvania, New Jersey, and Maryland), and ERCOT (Texas). Both regulated utility structure and deregulated wholesale market structure exist in different regions. While the two locations of Mountain View and Houston are in the deregulated electricity market regions, the other one in interests (i.e. Atlanta) is located in the regulated utility region (with fixed electricity price). For the two deregulated market regions, California has an hour-ahead forward market, whereas Texas has a 15-min ahead forward market. From the raw data we calculate the average daily electricity price for all these locations on May, 2009 and present the results in Figure 4 (www.caiso.com, www.ercot.com) [16]. We also calculate the hourly electricity price for all these locations on May 2nd, 2009 and present the results in Figure 5 [16]. Figure 6 [16] illustrates the different dynamics for Houston. Compared to the hourly market, the 15 mins real-time market is more volatile, with more high-frequency variations.

### 5.2 Performance Evaluation Based on Real-life Electricity Price Data for Google IDCs

Service providers like Google requires large computational resources in order to provide reasonable service. Google has a large number of data centers scattered around the world. At least 12 significant Google IDC installations are located in the United States. In this paper, we evaluate the performance of the proposed method with the electricity prices for three of these locations: Mountain View, California, Hous-

ton, Texas, and Atlanta, Georgia [32] [33]; and five front-end web portal servers.

To emulate Google IDC server power consumption, we assume that the total workload from five front-end web portal servers to three back-end IDCs in the three respective locations (i.e.  $C = 5$ ,  $N = 3$ ) is  $10^5$  requests per second; the delay constraint is 1 ms. We assume that the servers at one location is homogeneous, i.e.  $A_i$  and  $B_i$  are the same for each server at location  $i$ .  $A_i$  is a coefficient for calculating the effect of frequency changes and  $B_i$  represents the fixed energy consumption regardless of the current frequency setting. To derive server parameters, experimental data from [19, 29] are referred. Servers at all the locations are single processor machines: location 1 server has an Intel Pentium D 950 3.4 GHz CPU; location 2 server has an AMD Athlon 64 X2 4800+ 3.0 GHz CPU; location 3 server has an Intel Pentium 4 630 3.0 GHz CPU. In the simulation, we use  $p = 3$ . The parameters are summarized in Table 2, and Table 3.

**Table 2: Parameters of Three Google IDC Locations**

$i$	$f_{i,max}$	$M_i$	$D_i$	$A_i$	$B_i$
1	3.4	50000	0.001	3.206	68
2	2.4	50000	0.001	4.485	53
3	3.0	50000	0.001	2.370	70

**Table 3: Parameters of Five Google front-end web portal servers**

$j$	1	2	3	4	5
$L_j$	30000	15000	15000	20000	20000

To compare the total electricity cost of our proposed total electricity cost minimization method and the optimal power control method(i.e, without consideration of variable electricity price), we also solve the following constrained optimization problem:

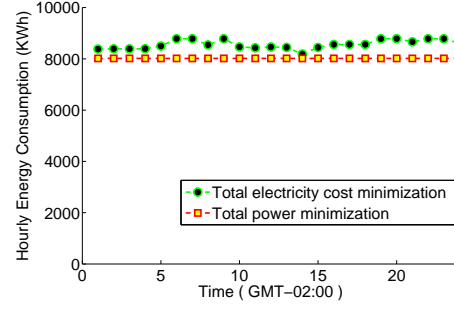
$$\min_{m_i, f_i} \sum_{i=1}^N m_i (A_i f_i^p + B_i),$$

subject to

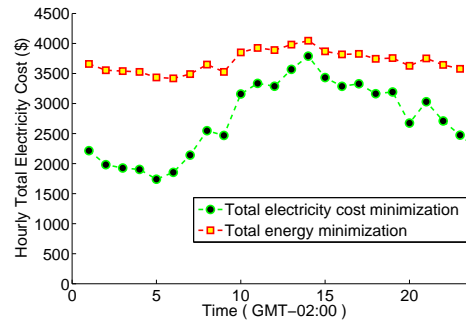
$$\begin{aligned} \sum_{i=1}^N \left( \frac{m_i f_i}{K} - \frac{1}{D_i} \right) &= L, \\ m_i &\leq M_i, \quad \forall i = 1, \dots, N, \\ m_i &\in \mathcal{N}, \quad \forall i = 1, \dots, N. \end{aligned}$$

Figure 7 shows the comparison of hourly total energy consumption for our proposed total electricity cost minimization method and the optimal power management while Figure 8 shows the comparison of hourly total electricity cost. As we can see, the optimal power management guarantees the lowest energy consumption all the time. However, lower power consumption does not ensure lower electricity cost in a multi-electricity-market environment. With total electricity cost minimization, the electricity cost is greatly reduced in every single hour, which indicates that the proposed method would help service providers save a lot in their electricity bill.

### 5.3 Server-Level Power Control at IDC locations



**Figure 7: Hourly Energy Consumption for Three Major Locations of Google IDCs on May 2nd, 2009**



**Figure 8: Hourly Total Electricity Cost for Three Major Locations of Google IDCs on May 2nd, 2009**

Figure 9 and Figure 10 show the CPU frequencies and the number of servers for each location to run to meet the workload requirement and the quality of service constraints for all the front-end web portal servers. First we observe that it is obvious that the CPU frequency and the number of servers at each location vary a lot with the electricity price variations, which indicates the optimal workload assignment changes with time for different locations. As a result, it is necessary to calculate the CPU frequency and the number of servers for different time at different locations. However, as the workloads at different locations may vary, there might be increasing bandwidth cost. In this paper, though we focus on the simple setup without considering bandwidth constraints, it is worth noting that our formulation can be easily extended to accommodate the bandwidth constraints. On the other hand, a service provider as large as Google can negotiate contracts with bandwidth providers on a nationwide basis.

### 5.4 Center-Level Load Balancing for IDCs

Figure 11 illustrates the optimal workload assignment for all of the three IDC locations. We can see that there is a coherent relationship between the variations of the hourly workload assignment shown in Figure 11 and the dynamics of the hourly electricity price shown in Figure 5. This indicates that the electricity prices at different IDC locations

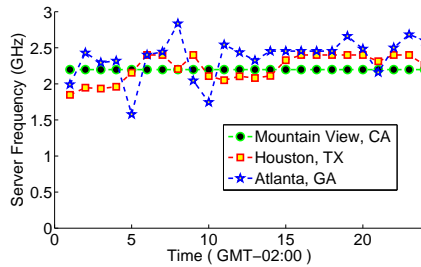


Figure 9: Server Frequency for Three Major Locations of Google IDCs on May 2nd, 2009

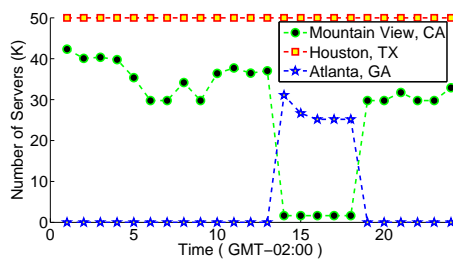


Figure 10: Number of Servers for Three Major Locations of Google IDCs on May 2nd, 2009

affect the amount of optimal assigned workload.

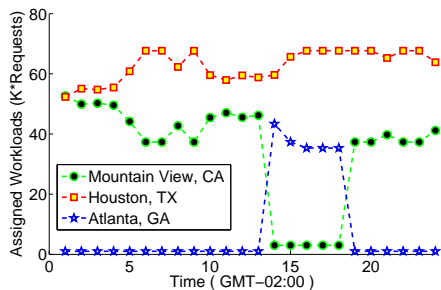


Figure 11: Assigned Workloads for Three Major Locations of Google IDCs on May 2nd, 2009

## 6. CONCLUSION

Power expenses are becoming an increasingly important fraction of the Internet Data Center (IDC) operating cost. In this paper, we investigate an emergent yet important problem of minimizing the total electricity cost for IDC under a multi-electricity-market environment, where the price of electricity could exhibit time and location diversities. We propose MEC-IDC, a joint load balancing and power management scheme that can minimize the total electricity cost for such an environment while guaranteeing the quality of

service to the end users. Extensive evaluations based on real-life electricity price data for multiple Google IDC locations demonstrates the effectiveness of MEC-IDC.

## Acknowledgment

The preliminary version of this work was done when Lei Rao and Xue Liu were in the School of Computer Science at McGill University. This work was supported in part by NSERC Discovery Grant 341823-07, NSERC Strategic Grant STPGP 364910-08, and FQRNT Grant 2010-NC-131844.

## 7. REFERENCES

- [1] Amazon Web Services LLC. Amazon elastic computing cloud (amazon ec2), 2008. [Online]. Available: <http://aws.amazon.com/ec2/>
- [2] A. M. Geoffrion. Generalized benders decomposition. *Journal of Optimization Theory and Applications*, 10(4):237-260, 1972.
- [3] A. Neumaier and O. Shcherbina. Safe bounds in linear and mixed-integer linear programming. *Mathematical Programming*, vol. 99, Mar. 2004, pp. 283-296.
- [4] A. Qureshi, R. Weber, H. Balakrishnan, J. Gutttag, and B. Maggs. Cutting the Electric Bill for Internet-Scale Systems. *ACM SIGCOMM*, 2009.
- [5] C. Courcoubetis and R. Weber. Pricing communication networks: economics, technology, and modelling. John Wiley & Sons Inc, 2003.
- [6] C. Hua, S. Wei and R. Zheng. Robust Channel Assignment for Link-level Resource Provisioning in Multi-radio Multi-channel Wireless Networks. In *Proceedings of the 16th IEEE International Conference on Network Protocols (ICNP)*, Orlando, FL, 2008.
- [7] E.A. Lee. Computing needs time. *Commun. ACM*, vol. 52, 2009, pp. 70-79.
- [8] G. Chen, W. He, J. Liu, S. Nath, L. Rigas, L. Xiao, F. Zhao. Energy-Aware Server Provisioning and Load Dispatching for Connection-Intensive Internet Services. In *Proceedings of the 5th USENIX Symposium on Networked Systems Design & Implementation (NSDI'08)*, San Francisco, CA, April 2008.
- [9] F. D. Galiana. An application of system identification and state prediction to electric load modeling and forecasting. *Electric Power Systems Laboratory, MIT, Cambridge, MA*, 1971.
- [10] Google. Google apps, 2008. [Online]. Available: <http://www.google.com/apps/intl/en/business/index.html>
- [11] H.W. Kuhn, P.J. Haas, I.F. Ilyas, G.M. Lohman, and V. Markl. The Hungarian method for the assignment problem. *Masthead*, vol. 23, 1993, pp. 1511C210.
- [12] Hewlett-Packard Development Company. HP Enterprise Power Calculators. <http://h30099.www3.hp.com/configurator/powercalcs.asp>, 2006.
- [13] J. Heo, D. Henriksson, X. Liu, T. Abdelzaher. Integrating Adaptive Components: An Emerging Challenge in Performance-Adaptive Systems and a Server Farm Case-Study. In *Proceedings of the 28th IEEE Real-Time Systems Symposium (RTSS'07)*, Tucson, Arizona, 2007

- [14] J. Liu, F. Zhao, X. Liu, W. He. Challenges Towards Elastic Power Management in Internet Data Centers. In Proceedings of 29th IEEE International Conference on Distributed Computing Systems Workshops(ICDCSW2009), pp.65-72, 2009.
- [15] L. A. Barroso and U. Högl. The case for energyproportional computing. *IEEE Computer*, vol. 40, no. 12, pp. 33-37, 2007.
- [16] 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 Proceedings of the 29th Annual IEEE Conference on Computer Communications (INFOCOM 2010), San Diego, USA, March 2010.
- [17] L. Sha, S. Gopalakrishnan, X. Liu, and Q. Wang. Cyber-Physical Systems: A New Frontier. Sensor Networks, Ubiquitous and Trustworthy Computing, 2008. SUTC '08. IEEE International Conference on, 2008, pp. 1-9.
- [18] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. Above the clouds: A Berkeley view of cloud computing. EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-28, Feb 2009. [Online]. Available: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-28.html>
- [19] M. Chin. Desktop CPU power survey. *Silentpcreview.com*, Apr. 2006.
- [20] Microsoft. Azure service platform, 2008. [Online]. Available: <http://www.microsoft.com/azure/default.msp>
- [21] M. Elnozahy, M. Kistler, R. Rajamony. Energy-Efficient Server Clusters. *Power-Aware Computer Systems*, 2003, pp. 179-197.
- [22] National Science Foundation. Cyber-physical systems. Technical report, NSF Workshop on Cyber-Physical Systems, 2006. <http://varma.ece.cmu.edu/cps/>.
- [23] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, A. Warfield. Xen and the art of virtualization. *ACM SIGOPS Operating Systems Review*, Volume 37 , Issue 5, December 2003.
- [24] R. Dorfman, P.A. Samuelson, and R.M. Solow. *Linear programming and economic analysis*. Courier Dover Publications, 1987.
- [25] R. Raghavendra, P. Ranganathan, V. Talwar, Z. Wang, and X. Zhu. No  $\alpha$ power struggles: coordinated multi-level power management for the data center. In ASPLOS XIII: Proceedings of the 13th international conference on Architectural support for programming languages and operating systems. New York, NY, USA: ACM, 2008, pp. 48-59.
- [26] R. Nathuji, K. Schwan. VirtualPower: coordinated power management in virtualized enterprise systems. Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles, Stevenson, Washington, USA, 2007.
- [27] S. Nedeveschi, L. Popa, G. Iannaccone, S. Ratnasamy, D. Wetherall. Reducing Network Energy Consumption via Sleeping and Rate-Adaptation. In Proceedings of the 5th USENIX Symposium on Networked Systems Design & Implementations (NSDI'08), San Francisco, CA, April 2008.
- [28] P. Skantze, M.D. Ilić and J. Chapman. Stochastic modeling of electric power prices in a multi-market environment. *IEEE Power Engineering Society Winter Meeting*, 2000.
- [29] Tom's Hardware. CPU performance charts. 2006.
- [30] T. Horvath and K. Skadron. Multi-mode energy management for multi-tier server clusters. Proceedings of the 17th international conference on Parallel architectures and compilation techniques, Toronto, Ontario, Canada: ACM, 2008, pp. 270-279.
- [31] T. Horvath, T. Abdelzaher, K. Skadron, and X. Liu. Dynamic voltage scaling in multitier web servers with end-to-end delay control. *Computers, IEEE Transactions on*, vol. 56, no. 4, pp. 444-458, April 2007.
- [32] United States Energy Information Administration, Department of Energy, available online at [www.eia.doe.gov](http://www.eia.doe.gov).
- [33] United States Federal Energy Regulatory Commission, available online at [www.ferc.gov](http://www.ferc.gov).
- [34] W. Wolf. Cyber-physical Systems. *Computer*, vol. 42, 2009, pp. 88-89.
- [35] X. Liu, J. Heo, L. Sha, X. Zhu. Adaptive Control of Multi-Tiered Web Application Using Queueing Predictor. In Proceedings of the 10th IEEE/IFIP Network Operations and Management Symposium (NOMS 2006), Vancouver, Canada, 2006.
- [36] X. Fan, W. Weber, L. A. Barroso. Power provisioning for a warehouse-sized computer. In Proceedings of the 34th annual international symposium on Computer architecture, San Diego, California, USA, 2007.