Data Center Power Cost Optimization Via Workload Modulation

Cheng Wang, Bhuvan Uргаonkar, Qian Wang, George Kesidis, and Anand Sivasubramaniam
Departments of CSE, EE, and MNE
The Pennsylvania State University

Abstract—We formulate optimization problems to study how data centers might modulate their power demands for cost-effective operation taking into account various complexities exhibited by real-world electricity pricing schemes. For computational tractability reasons, we work with a fluid model for power demands which we imagine can be modulated using two abstract knobs of demand dropping and demand delaying (each with its associated penalties or costs). We consider both stochastically known and completely unknown inputs, which are likely to capture different data center scenarios. Using empirical evaluation with both real-world and synthetic power demands and real-world prices, we demonstrate the efficacy of our techniques.

I. INTRODUCTION

It is well-known that data centers incur significant costs towards powering and cooling their computing, networking, and storage (IT) equipment, making the optimization of these costs an important research area. Most electric utilities employ one (or a combination) of the following pricing mechanisms: (i) *peak-based pricing*, wherein a component of the electricity bill is dependent on the peak power drawn over the billing cycle (typically a month) [2], (ii) *time-varying pricing*, wherein the per unit energy price fluctuates over time (there are many examples of this operating at different time granularity such as very fine time-scale spot prices [11] or higher prices during periods of “coincident peaks” experienced by the utility [7]), and (iii) *load-dependent or tiered pricing*, wherein higher energy prices are applied for higher energy consumption [9].

Whereas recent research has begun looking at data center cost optimization bearing some of these pricing complexities in mind, this is still a nascent field in many ways. In particular, most instances of such work (e.g., [4]) focus exclusively on a particular pricing feature, whereas real-world tariffs often combine these features, making cost-effective data center operation even more complicated. This is the context of our paper: *how should data centers optimize their costs given the various features of real-world electricity pricing?*

Broadly speaking, the vast literature on data center power cost optimization may be understood as employing one or both of (i) demand-side (based on modulating the data center’s power demand from within) and/or (ii) supply-side (based on employing additional sources of energy supply the data center’s existing utility provider and backup sources) techniques. Our focus is on techniques based on “IT knobs,” a subset of (i) that relies

upon software/hardware mechanisms to change the power consumption within the data center. Other demand-side modulation based on local generation capabilities or energy storage [4] and a large body of work that has emerged on supply-side [10] are complementary to our work, and cost-effective operation using these techniques is interesting future work.

We make the following research contributions:

- *Problem Formulation*: We formulate an optimization problem for a data center that wishes to employ IT knobs for power demand modulation for cost-effective operation. Our key novelty over related work in this area is our incorporation of various features of real-world electricity pricing schemes into a single unified formulation.
- *Algorithm Design*: Given that power demands and electricity prices can exhibit uncertainty, we devise a stochastic dynamic program (SDP) that leverages predictive workload and price models. We also devise approximations for our SDP that might be useful for scenarios where the exact SDP proves computationally intractable. For scenarios with poor input predictability, we also explore fully online algorithms, and prove a competitive ratio of 2 for one of them.
- *Empirical Analysis*: We evaluate the efficacy of our algorithms in offering cost-effective operation for a variety of workloads derived from real-world traces. We find that our stochastic control techniques offer close-to-optimal cost savings when the workloads exhibits strong time-of-day patterns; our online algorithms performs well even when there is unpredictable flash crowd in the workload.

II. PROBLEM FORMULATION

A. Context and Key Assumptions

We consider our problem at the granularity of an entire data center or a “virtual” data center (i.e., a subset of the data center whose IT resources are dedicated to a “tenant” application that is allowed to carry out its own power cost optimization). For reasons of computational feasibility, we make two key assumptions/simplifications. First, we imagine power demands as being “fluid,” when, in practice, power cost optimization must deal with discrete resource allocations and software job characteristics. Second, we choose to capture the IT power/performance/cost trade-offs via two abstract knobs: demand dropping (e.g., a

Web search application with partial execution for meeting response time targets) and demand delaying (e.g., some MapReduce-based batch workloads can be delayed as long as it finishes before a deadline). We assume that we can leverage existing or future work on translation mechanisms between our fluid power demand and its modulation via delaying/dropping (on the one hand), and actual IT resources, their control knobs, and software characteristics and performance (on the other hand). Some examples of such translations may be found in [12].

We employ convex non-decreasing functions to model the loss due to demand delaying ($l_{\text{delay}}(\text{demand}, \text{delay})$) and dropping ($l_{\text{drop}}(\text{demand})$), and denote by τ the delay tolerance for a unit of delayed power demand.

B. Notation

Input Parameters: We consider a discrete-time model wherein control decisions are made at the beginning of *control windows* of equal duration δ . We denote by T the number of such control windows within a single billing cycle, which constitutes our *optimization horizon*. Let $\{p_t : 1 \leq t \leq T\}$ denote the power demand of the (possibly virtual) data center over the optimization window. We define a time-varying energy price of $\alpha_t \$/kWh$ and $\beta \$/kW$ to represent the time-varying pricing and peak-based charging in a billing cycle, respectively.

Decision Variables: At the beginning of control window t (or simply t henceforth), the data center can have “residual” demands from the past (in addition to the newly incoming demand p_t). Given that demands may not be delayed by more than τ windows, these residual demands may have originated in (i.e., first postponed during) the set of windows $h(t) = \{\max\{1, t - \tau\}, \dots, t - 1\}$. We denote these residual demands as the set $\{r_{i,t}\}_{i \in h(t)}$. We denote the peak demand admitted in any window during $[1, \dots, T]$ as $y_{\max}$. The control actions to be taken by the data center during t involve admitting, postponing, and dropping portions of $r_{i,t}$ and p_t . We denote as $a_{i,t}$ and $d_{i,t}$ ($i \in h(t)$) the portion of residual demand $r_{i,t}$ that are admitted and dropped during t ; as $a_{t,t}$ and $d_{t,t}$ the portion of p_t that are admitted and dropped during t (i.e., without incurring any delay). We denote as $h^+(t)$ the set $h(t) \cup \{t\}$.

C. Offline Decision Making

Objective: We choose as our data center’s objective the minimization of the sum of its utility bill and any revenue loss resulting from demand modulation:

$$\mathcal{O}(\{\mathcal{A}_t\}, \{\mathcal{D}_t\}) \stackrel{\text{def}}{=} \beta y_{\max} + \sum_t \left(\alpha_t \sum_{i \in h^+(t)} a_{i,t} + \sum_{i \in h^+(t)} l_{\text{drop}}(d_{i,t}) + \sum_{i \in h(t)} l_{\text{delay}}(a_{i,t}, t - i) \right).$$

Constraints: The aggregate demand in the data center at the beginning of t is given by the new demand p_t , and any demand unmet so far (deferred from previous time slots $h(t)$) $\mathcal{R}_t$. Since this demand must be treated via a combination of the following three: (i) serve demand

$a_{t,t}$, (ii) drop demand $d_{t,t}$, and (iii) postpone/delay demand ($r_{t,t+1}$) (to be served during $[t + 1, \dots, T]$), we have:

$$p_t - a_{t,t} - d_{t,t} = r_{t,t+1}, \quad \forall t. \quad (1)$$

The residual demand from $i \in h(t)$ that is not admitted during t is either dropped during t or postponed to the next time slot $t + 1$:

$$r_{i,t} - a_{i,t} - d_{i,t} = r_{i,t+1}, \quad i \in h(t), \quad \forall t. \quad (2)$$

Any demand that has been postponed for τ time slots may not be postponed any further:

$$r_{t-\tau,t} - a_{t-\tau,t} - d_{t-\tau,t} = 0, \quad \forall t. \quad (3)$$

To keep our problem restricted to one billing cycle, we add an additional constraint that any delayed demand (even if it has been postponed for less than τ time slots) must be admitted by the end of our optimization horizon:

$$r_{i,T+1} = 0, \quad i \in h(t). \quad (4)$$

Alternate formulations that minimize costs over multiple billing cycles may relax the constraint above. The peak demand admitted $y_{\max}$ must satisfy the following:

$$y_{\max} \geq \sum_{i \in h^+(t)} a_{i,t}, \quad \forall t. \quad (5)$$

Finally, we have:

$$y_{\max}, a_{i,t}, d_{i,t}, r_{i,t} \geq 0, \quad i \in h^+(t), \quad \forall t. \quad (6)$$

Offline Problem: Based on the above, we formulate the following offline problem (called OFF):

$$\begin{aligned} &\text{Minimize} && \mathcal{O}(\{\mathcal{A}_t\}, \{\mathcal{D}_t\}), \\ &\text{subject to} && (1) - (6). \end{aligned}$$

III. STOCHASTIC CONTROL

Whereas OFF can be useful in devising cost-effective workload modulation (as we will corroborate in Section V), many data center workloads exhibit uncertainty. Deviations in demands or prices compared to those assumed by OFF can result in poor decision-making. Consequently, it is desirable to devise workload modulation techniques that can adapt their behavior to workload and price evolution in an *online* manner.

A. Stochastic Dynamic Program

Data center workloads can often be captured well via statistical modeling techniques. Therefore, before devising fully online algorithms in Section IV, we first develop an optimization framework based on stochastic dynamic programming (SDP) that leverages predictive models for workload demands and electricity prices.

Computational Tractability Concerns: Unfortunately, in its general form, our SDP is likely to be computationally prohibitive due to the well-known “curse of dimensionality”. If we denote as $O(R)$ the run-time of sub-problem

of SDP, as L_p and L_α the number of discretization levels used for power demands and energy price, respectively, under stage-independence assumptions on demands and prices, the runtime for our SDP can be expressed as $O(R \cdot L_p^{2(\tau+2)} \cdot L_\alpha \cdot T)$.

B. Scalable Approximations of Our SDP

We investigate two successive simplifications to our SDP for overcoming its scalability problems caused by large values of τ .

Linear Approximation of Delay Penalties: Our first approximation, called SDP_{Lin} , employs a linear approximation $l_{\text{delayLin}}(\cdot)$ for the function $l_{\text{delay}}(\cdot)$. Specifically, if $l_{\text{delayLin}}(x, y)$ has the form $k_{\text{delay}}xy$ for $k_{\text{delay}} > 0$, then it is easily seen that the SDP state can be simplified to the 2-tuple $s_t = (y_t, r_t)$, where r_t denotes the sum of the residual demands ($r_{i,t}$) at the beginning of t . SDP_{Lin} offers a significantly reduced runtime of $O(R \cdot L_p^5 \cdot L_\alpha \cdot T)$. Of course, this is at the cost of poorer solution quality, and we empirically evaluate this trade-off in Section V.

Ignoring the Delay Knob: Our second approximation, called SDP_{Drop} , is based on completely ignoring the (computationally) problematic knob of demand delaying. Notice that it is a special case of SDP_{Lin} in that it can be viewed as employing $\tau = 0$. SDP_{Drop} has a reduced runtime of $O(R \cdot L_p^3 \cdot L_\alpha \cdot T)$.

IV. ONLINE ALGORITHMS

We devise two online control policies that might prove useful in scenarios where predictive models for power demands fail.¹ Due to space constraints, we leave the details of the two online algorithms in [12].

ON_{Drop} : Online Algorithm With Only Dropping: Many workloads are delay-intolerant but do allow (possibly implicit) dropping of the power demands they pose with associated loss in revenue. As two simplifications, we assume (i) $l_{\text{drop}}(x)$ has the form $k_{\text{drop}}x$ as in Section V, and (ii) $\alpha_t = \alpha, \forall t$. We leave the extension of ON_{Drop} and its competitive analysis without these assumptions for future work.

The following lemma is key to the design of ON_{Drop} .

Lemma 1. *If only demand dropping is allowed, $l_{\text{drop}}(x) = k_{\text{drop}}x$, and $\alpha_t = \alpha$, then the optimal demand dropping threshold θ has the following form: if we denote $\hat{p}_t$ as the t^{th} largest demand value in $\{p_t\}_{t=1}^T$, then $\theta = \hat{p}_n$, where $n = \lceil \frac{\beta}{(k_{\text{drop}} - \alpha)} \rceil$.*

We exploit Lemma 1, to devise ON_{Drop} as follows:

- 1) Initialization: Set demand dropping threshold $\theta = 0$, $n = \lceil \frac{\beta}{(k_{\text{drop}} - \alpha)} \rceil$.
- 2) At time t , sort $p_1, \dots, p_t$ into $\hat{p}_1, \dots, \hat{p}_t$ such that $\hat{p}_1 \geq \hat{p}_2 \geq \dots \geq \hat{p}_t$.
- 3) Update θ as follows: If $t < n$, $\theta = 0$; Otherwise, $\theta = \hat{p}_n$.

¹Extending these algorithms to also deal with uncertainty in electricity prices is an important direction for future work.

- 4) Decision-making: Admit $\min(p_t, \theta)$, drop $[p_t - \theta]^+$.

Theorem 1. ON_{Drop} offers a competitive ratio of $2 - \frac{1}{n}$.

We present the proof in [12].

ON_{MPC} : Model Predictive Control: Finally, we also develop an online algorithm by adapting ideas from Model Predictive Control (MPC), a well-known suboptimal control theory. Our adaptation is simple and assumes perfect input knowledge for a small lookahead window h , using which our algorithm solves a smaller version of OFF over an optimization horizon $H \leq T$. Our purpose in devising this algorithm is to employ it as a baseline that we empirically compare with our other algorithms.

V. EXPERIMENTAL RESULTS

We evaluate our algorithms using both power demands derived from real-world traces and constructed synthetically. Our metric is "cost savings" which is the percentage reduction in costs due to a particular algorithm compared to doing no demand modulation. We assume $l_{\text{drop}}(x) = k_{\text{drop}}x$ and $l_{\text{delay}}(x, t) = k_{\text{delay}}t^2x$, respectively.

A. Parameters and Workloads

Param.	Description	Value
α, β	Energy price (\$/kWh), Peak power price (\$/kW/month)	0.046, 17.75 [11]
k_{drop}	Cost incurred by dropping demand (\$/kWh)	0.72
k_{delay}	Cost incurred per δ for delaying (\$/kW/month))	0.02 [6]
τ	Maximum delay allowed	6δ (1 hour)
δ	Control window	10 minutes
T	Optimization window	4320 (30 days / δ)
H, h	Receding horizon, Lookahead window	1 day 6 hours

TABLE I: Various problem parameters.

Table I lists the values used for various parameters in our evaluation. We employ three real-world power demands presented in recent studies: Google, Facebook, MediaServer (streaming media) [1], [3], [5]. Additionally, we create a synthetic power demand series with an emphasis on including an unpredictable surge in power demand. Synthetic is built by adding a high power surge to the demand for MediaServer on the 15th day.

B. Peak Pricing Schemes

Table II presents the cost savings (%) offered by different algorithms under the peak-based pricing scheme with the parameters (α, β) shown in Table I.

Key Insights: (i) For our workloads and parameters, delaying demand offers larger benefits than dropping, (ii) workload properties strongly affect the cost savings under peak-based tariff, and (iii) our stochastic control techniques are able to leverage workload prediction to offer near-optimal cost savings when there is no unexpected flash crowd.

Control knobs	Only dropping				Only delaying			Dropping+Delaying		
Workload	OFF	ON _{MPC}	SDP _{Drop}	ON _{Drop}	OFF	ON _{MPC}	SDP _{Lin}	OFF	ON _{MPC}	SDP _{Lin}
Google	1.45	-0.21	1.14	-0.29	10.74	9.92	2.83	10.74	9.92	2.83
Facebook	13.11	9.26	12.52	8.95	16.39	9.96	11.79	16.97	9.96	11.79
MediaServer	13.82	11.21	8.13	11.66	27.64	27.16	18.26	27.64	27.16	18.26
Synthetic	36.56	30.80	29.86	32.49	38.42	32.40	-10.37	45.05	32.43	-10.37

TABLE II: Cost savings (%) offered by different algorithms under peak-based tariff.

Fig. 1: Workloads from the 14th day to the 16th day. The blue line is the original power demand, and the red line is the time-of-day behavior of the demand.

C. Time-varying Prices

We assume dropping cost is always larger than energy cost ($l_{\text{drop}}(x) > \alpha_t x, \forall t$), thus the only control knob we can choose is deferring demand. We set $k_{\text{delay}} = 0.01\$/kWh$ as in [6], and evaluate cost savings both for linear delay costs ($l_{\text{delay}}(x, t) = k_{\text{delay}}tx$) and quadratic delay costs ($l_{\text{delay}}(x, t) = k_{\text{delay}}t^2x$). In our experiments, we use the time-varying price time-series from [8], which is charged on commercial or industrial customers ($> 2MW$).

Control knob	Only delaying	
Workload	Linear delay cost	Quadratic delay cost
Google	9.21	7.97
Facebook	3.02	1.55
MediaServer	7.95	6.45
Synthetic	7.77	6.34

TABLE III: Cost savings (%) of different delay cost model under time-varying pricing tariff.

Our observations are presented in Table III. **Key Insights:** (i) workload properties appear to have less impact on cost savings compared to peak-based pricing, and (ii) cost savings greatly depend on the energy price fluctuation and the actual delay penalty (linear vs. quadratic).

VI. CONCLUSIONS AND FUTURE DIRECTIONS

We formulated optimization problems to study how data centers might modulate their power demands for cost-effective operation given various complexities exhibited by

real-world electricity pricing schemes. For computational tractability reasons, we worked with a fluid model for power demands which we imagined could be modulated using two abstract knobs of demand dropping and demand delaying. We consider both stochastically known and completely unknown inputs, which are likely to capture different data center scenarios. We demonstrated the efficacy of our techniques using empirical evaluation with both real-world and synthetic power demands and real-world prices.

VII. ACKNOWLEDGEMENT

This research was supported, in part, by the following: NSF CAREER award CNS-0953541, NSF CCF-1228717, NSF 1000699, NSF 1302225 and NSF 1205618.

REFERENCES

- [1] Y. Chen, A. Ganapathi, R. Griffith, and R. H. Katz. The case for evaluating mapreduce performance using workload suites. In *MASCOTS*, pages 390–399, 2011.
- [2] Duke utility bill tariff, 2012. <http://www.considerthecarolinas.com/pdfs/scscheduleopt.pdf>.
- [3] X. Fan, W. Weber, and L. A. Barroso. Power provisioning for a warehouse-sized computer. In *ISCA*, pages 13–23, 2007.
- [4] S. Govindan, D. Wang, A. Sivasubramaniam, and B. Urgaonkar. Leveraging stored energy for handling power emergencies in aggressively provisioned datacenters. In *ACM ASPLOS*, pages 75–86, Mar 2012.
- [5] B. Li, G. Y. Keung, S. Xie, F. Liu, Y. Sun, and H. Yin. An empirical study of flash crowd dynamics in a p2p-based live video streaming system. In *GLOBECOM*, pages 1752–1756, 2008.
- [6] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. H. Andrew. Greening geographical load balancing. *SIGMETRICS '11*, pages 233–244, New York, NY, USA, 2011. ACM.
- [7] Z. Liu, A. Wierman, Y. Chen, and B. Razon. Data center demand response: Avoiding the coincident peak via workload shifting and local generation. In *ACM SIGMETRICS*, Jun 2013.
- [8] Large General TOU, National Grid, 2013. http://www.nationalgridus.com/niagaramohawk/business/rates/5_elec_sc3a.asp.
- [9] Ontario Energy, 2013. <http://www.ontarioenergyboard.ca/OEB/Consumers/Electricity/Electricity+Prices#tiered>.
- [10] C. Ren, D. Wang, B. Urgaonkar, and A. Sivasubramaniam. Carbon-aware energy capacity planning for datacenters. In *MASCOTS*, pages 391–400. IEEE Computer Society, 2012.
- [11] Time-of-use large general service, SCEG, 2013. <http://www.sceg.com/en/commercial-and-industrial/rates/electric-rates/default.html>.
- [12] C. Wang, B. Urgaonkar, Q. Wang, G. Kesidis, and A. Sivasubramaniam. Data center cost optimization via workload modulation under real-world electricity pricing. Technical report, Penn State University. Available: <http://www.cse.psu.edu/research/publications/tech-reports/2013/CSE-13-010.pdf>, 2013.