Data Center Optimal Regulation Service Reserve Provision with Explicit Modeling of Quality of Service Dynamics

Hao Chen[1], Bowen Zhang[2], Michael C. Caramanis[3], and Ayse K. Coskun[4]

Abstract— **Data centers have shown great opportunities to participate in extensive demand response programs in recently years. This paper specifically focuses on data centers as participants in regulation service reserves (RSR) power market. We propose a novel approach to model the dynamics of the job processing Quality of Service (QoS) in data centers that offer RSR, and use stochastic dynamic programming (DP) to solve for the optimal reserve deployment policies. We show that the job QoS degradation can be modeled as a time varying probability distribution function (PDF) whose mean and variance evolve as functions of recent control statistics. The mean and variance are in fact additional state variables or sufficient statistics of the stochastic DP whose solution provides the data center operator (DCO) decision supports to minimize the average operating costs associated with RSR signal tracking error and job processing QoS degradation. Simulation results show that the feedback control policy obtained from the stochastic DP solution can reduce the DCO's operating costs compared to heuristic operating protocols reported in the literature. In addition, the DP value function can assist the DCO to bid optimally into the hour-ahead joint energy and reserve market.**

I. INTRODUCTION

Power system renewable integration is increasing world wide. The EU has set the goal of reaching a 20% share of renewable energy in gross energy consumption by 2020 [1]. In the US, 38 states have long term renewable portfolio standards and 14 states have installed more than 1,000 MW of wind power [2]. It is expected that the total renewable generating capacity will have a growth of 52% utill 2040 [3]. Higher renewables integration increases the frequency control, regulating and operating reserve requirements provided mostly by conventional centralized generation today but expected to be significantly complemented in the near future by demand side participants, such as smart buildings, large factories, and data centers through either centralized utility direct load control (DLC) or price based distributed control.

An interesting related development is that electricity consumed by data centers is growing rapidly. In the US, the annual growth rate of data center electricity consumption is 12% [4], while it has already accounted for about 3% of overall electricity consumption [5]. Data centers have had a great impact on today's power market. Recent advanced server power management techniques, such as dynamic voltage frequency control (DVFS) [6] and management of CPU resource limits [7], have enabled data centers to use the flexibility in their power consumption to advantage. In this vein we argue that data centers offer a unique opportunity to participate in extensive demand response programs, either through current and emerging technology such as DLC, or as participants in reserve power markets on a par basis with centralized generation. Such a new role of data centers will render their growth economically and environmentally sustainable, benefiting data centers themselves and the electricity sector in general, by enabling efficient integration of clean energy generation.

This paper investigates the participation of data centers in the provision of regulation service reserves (RSR) ([8], [9]), by modulating the data center servers processing rates to meet system needs, while maintaining job processing Quality of Service (QoS) promised to their customers. We consider RSR provision in the hour-ahead market that clears with relatively high RSR prices. Our specific contributions are:

First, we introduce a dynamic model of probability distribution function (PDF) that quantifies the likelihood of a data center job departing beyond the contracted QoS guarantees (Section II-C). Based on extensive numerical study, the dynamic QoS degradation PDF can be accurately approximated by a uniform distribution whose mean and variance are functions of the history of the control inputs (i.e., server retired instructions per second (RIPS) as the processing rate, or equivalently power consumption rate). Specifically, the mean of the QoS degradation PDF can be estimated by a linear regression of the integral of recent controls, while the variance can be represented as a linear regression of the number of job departures, which can be characterized as a Poisson random variable whose parameters are related to the current control. We account for the fact that the protracted use of a high processing rate decreases the likelihood of large QoS degradation, and vice versa.

Second, we develop a stochastic dynamic programming (DP) problem to minimize the expected cost of tracking errors in RSR provision and QoS guarantee violations (Section II-C). The state variables of DP include the distribution of QoS degradation (i.e., the mean and variance of it, as they are sufficient statistics to characterize the distribution), RSR signal transmitted by the independent system operator (ISO), and statistics on the recent data center power consumption trajectory. The DP is solved using value iteration with

The authors gratefully acknowledge support of the U.S. National Science Foundation under EFRI Grant 1038230.

[1]Corresponding author. H. Chen is with the Department of ECE, Boston University, Boston, MA, 02215 USA e-mail: haoc@bu.edu.

[2]B. Zhang is with the Division of Systems Engineering, Boston University, Boston, MA, 02215 USA e-mail: bowenz@bu.edu.

[3]M. C. Caramanis is with the Department of Mechanical Engineering and the Division of Systems Engineering, Boston University, Boston, MA, 02215 USA e-mail: mcaraman@bu.edu.

[4]A. K. Coskun is with the Department of ECE, Boston University, Boston, MA, 02215 USA e-mail: acoskun@bu.edu.

Monte Carlo simulated state transitions to derive optimal state feedback policies and the average cost value function (Section III-B). We observe and discuss on the partial non-monotonic property in the derived optimal policies, which is caused by the discontinuity in the cost of job system time that quantifies contracted QoS degradation.

Finally, the DP based optimal policy is compared with heuristic operating policies proposed in previous work on data center RSR provision [10] (Section III-C). Results show that our control law is associated with costs that are 5% lower than those incurred under the heuristic policies. In addition, we show that the DP value function can assist the data center operator (DCO) to optimally bid into the hour ahead energy and reserve market (Section IV).

Broadly speaking, this work also contributes to the real time control literature by relaxing the usual assumption that the QoS, or consumer's utility preferences, can be adequately characterized by a static PDF in the short-term ([11], [12], [13], [14]). For example, the demand for energy in the cooling zone, or the demand for access to a dynamically priced mobile service bandwidth, has been commonly modeled by a uniform probability distribution that remains constant over time regardless of past controls (e.g., prices). Undoubtedly, this assumption is inaccurate and the PDF of consumer preferences can be affected significantly by the recent control [15].

II. PROBLEM FORMULATION

In this section, we first introduce the RSR market for data centers to participate. Then we present the general data center model as well as the applicable server power management techniques in RSR provision. After that, we discuss our dynamic QoS modeling, and introduce the formulation of the data center RSR provision as a stochastic DP problem.

A. Regulation Service Reserves

Today's power markets and reserve provisions are classified into several categories based on different time scales and the frequency of the reserve commands deployed ([9], [16], [17]). In this work, we focus on RSR in the hour-ahead power market, as data centers are capable of modulating their powers at such frequency (4 seconds), and the price of reserves is high. While previous RSRs were mainly provided by centralized generators, market rules are changing to encourage more demand sides to provide reserves. For instance, PJM has allowed demand sides to provide RSR since 2006 [18], and other ISOs are following this trend.

In RSR provision, each potential provider bids an average power consumption $\bar{P}$, and a reserve value R to the ISO in an hour ahead. Once the bid is accepted, and the market prices of energy consumed and reserve are cleared at Π^E and Π^R respectively, the RSR provider is charged for $\Pi^E\bar{P}-\Pi^R R$ in the following hour. In other words, the provider can receive $\Pi^R R$ credits for providing R amount of reserves. The credits, however, do not come for free. The RSR provider is asked to track the RSR signal $y(t)$ broadcast every 4 second, by modulating its power consumption $P(t)$ such that $P(t)\approx\bar{P}+y(t)R$. The signal $y(t)$ is the main tool used by ISO to balance the supply and demand in the power market. It is generated in real-time that is unknown to providers in the hour ahead. However, the statistical behavior of $y(t)$ is well known. It is a random variable between $[-1,1]$, with an average of zero over long time intervals. The signal is updated every 4 seconds in increments that do not exceed $\pm R/(\tau/4)$, where τ is 150 seconds for the fast signal and 300 seconds for the slow signal. It follows a well-behaved two level Markov model whose transition probabilities can be calibrated in advance. During the hour, the tracking error $|\varepsilon(t)|=|P(t)-(\bar{P}+y(t)R)|$ is also calibrated, and part of credits $\Pi^R R$ is reduced based on it. Examples of the signal as well as its detailed descriptions are referred to ([10], [19]).

B. The Data Center Model

A data center consists of many servers. Each server can be set in one of multiple states at any time: active, idle, sleep and shut down. Only in the active state a server has its processing rate larger than zero, hence, can serve jobs. When the server is active, the power consumption of it can be modulated by several power management techniques, such as DVFS [6] and CPU resource limits [7]. Meanwhile, the server processing rate is changed correspondingly. Previous experiments have shown that applying either DVFS or CPU resource limits as the control knob, the relation between server dynamic power consumption $P_s(t)$ and the processing rate $u_s(t)$ at time t can be well fitted by a linear function f_p [10]:

$$P_s(t)=f_p(u(t))=ku_s(t)+P_{idle}, \tag{1}$$

where P_{idle} is the power consumption of the server at the idle state, which is a constant[1], k is a fixed parameter depending on types of jobs that are served. In this work for simplicity, we mainly focus on two most common states of the server in emerging data centers – the active state and the idle state, while leaving rest of states to be considered in future.

A data center total power consumption is an aggregation of the consumption from several sub-components, such as computational (including all servers), cooling (air conditioners, fans, etc.) and lighting units. We focus on regulating the computational power in this work, while other units in general have lower regulation capacity, thus can participate in some slower frequency demand responses. The computational power is an aggregation of each server's consumption, i.e., $P(t)=\sum_{s=1}^{N}P_s(t)$ at any time t, where $P(t)$ is the total power consumption of the computational units, $P_s(t)$ is each server's power (we use subscript s to denote server level variables in this paper), and N is the number of servers in the data center. In addition, we assume each server can only serve one job per time, and define a queue for holding incoming jobs of the data center.

C. The Stochastic Dynamic Programming Problem

We formulate the data center RSR provision as a stochastic DP problem. We start with the scenario when all servers are

[1] P_{idle} in fact is a temperature dependent variable. We assume there is no temperature change here.

in active state and have the same controllable processing rate $u(t)$ at time t, i.e., the total data center power budget $P(t)$ is always uniformly distributed to each server, so that the fairness among all servers is kept. The period cost function of the DP is composed of (i) the cost of inaccurate RSR signal tracking, i.e., $PC_{track}(t)$, characterized by the deviation between the data center power consumption $P(t)$ and the RSR signal $y(t)$, and (ii) the cost of QoS degradation, i.e., $PC_{QoSD}(t)$, characterized by the PDF of QoS degradation. In later discussion, we introduce to represent the PDF of QoS degradation by its mean and variance.

(i) Tracking cost $PC_{track}(t)$: denoting the processing rate of an individual server by $u_s(t)$, it has been shown that the server power consumption $P_s(t)$ is linearly related to $u_s(t)$ with function f_p in Eq.(1). Since all servers operate at the same controllable processing rate $u(t)$ for fairness, the whole data center energy consumption is $Nf_p(u(t))$. Given the RSR signal $y(t)$, the tracking error period cost is defined as:

$$PC_{track}(t) = \Pi^{Err}|Nf_p(u(t)) - (\bar{P} + y(t)R)|, \tag{2}$$

where Π^{Err} is a constant, representing the penalty price on per unit of tracking error.

(ii) Cost of QoS degradation $PC_{QoSD}(t)$: we first define the QoS degradation of each job i as $QoSD_i = T_i/T_{i,\min}$, where T_i is the job system time (i.e., waiting time plus processing time), $T_{i,min}$ is the job shortest processing time, which is a static value measured in advance, referring to the time of processing the job with the maximal server processing rate $u_{\max}$, and without any waiting time in the queue. $QoSD_i = 1$ means no degradation in QoS of job i.

We start with simulating a data center with $N = 1000$ servers extensively to characterize the distribution of the dynamic QoS degradation for each 4 seconds[2]. Assuming each server's maximal possible processing rate is $u_{\max}$, we simulate the scenario that jobs arrive following a Poisson distribution with the parameter $\lambda = 50\% * Nu_{\max}$, where $Nu_{\max}$ represents the maximal processing capacity of the data center. A job arrival rate at 50% of maximal capacity is selected here for the reason that an utilization around 50% is a typical scenario in emerging data centers.

Based on queueing theory, in order to guarantee that the system is stable, the average processing rate of the whole data center, i.e., $N\bar{u}$, should be greater than the job arrival rate $\lambda = N * 0.5u_{\max}$. Hence the constraints on the data center RSR bidding values $(\bar{P}, R)$, where $\bar{P} = Nf_p(\bar{u})$, are as follows:

$$\begin{aligned} &Nf_p(u_{\max}) \geq \bar{P} > Nf_p(0.5u_{\max}), \\ &\min(Nf_p(u_{\max}) - \bar{P}, \bar{P} - NP_{idle}) \geq R \geq 0. \end{aligned} \tag{3}$$

We simulate by using a 24-hour historical PJM RSR signal data [8] as $y(t)$, and test on different selections of $(\bar{P}, R)$ that satisfy Eq.(3). In addition, since different control policies lead to varying tracking errors, for the general purpose we involve the tracking error $\varepsilon(t) = Nf_p(u(t)) - (\bar{P} + y(t)R)$ as a Gaussian random variable in simulation,

[2]As mentioned in Section II-A, 4 seconds is the frequency of the RSR signal regulated. For the rest of paper, by default the time interval t is 4 seconds.

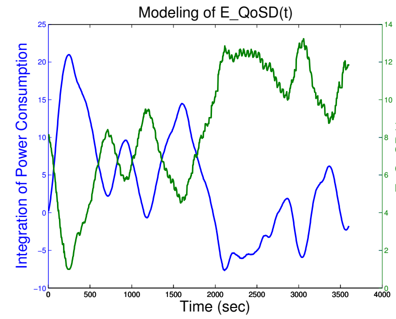

Fig. 1. Mean of the QoS degradation is characterized by the integration of the past power consumption. Strong anti-correlation -0.97 is found between two curves.

i.e., $\varepsilon(t) \sim N(\mu_\varepsilon(t), \sigma_\varepsilon^2(t))$, where $\mu_\varepsilon(t)$ is changed for every $t = 5$ minutes, obeying a uniform distribution as $\mu_\varepsilon(t) \sim U(-0.1R, 0.1R)$, and $\sigma_\varepsilon^2(t) = |4\mu_\varepsilon(t)|$. During the simulation, jobs are served with processing rate $u(t)$ that calculated based on the assigned data center power budget $\bar{P} + y(t)R + \varepsilon(t)$. We record the QoS degradation of every job that departs the system in the 4 second interval to generate the distribution of QoS degradation for every 4 seconds.

Simulation results show that in every 4 seconds the QoS degradation is uniformly distributed. Therefore it is necessary and sufficient to characterize the PDF by its mean and variance. We begin by formulating the mean of QoS degradation E_QoSD(t) for every 4 seconds. Based on standard queuing theory, the mean of QoS degradation depends only on the mean of queuing length μ_W of the system, when the data center does not provide RSR and consumes power steadily at level $\bar{P}$. When providing RSR, if the power consumption $Nf_p(u(t))$ is higher than $\bar{P}$, then it results in a smaller value of μ_W and therefore smaller E_QoSD(t), and vice versa. We further observe a strong anti-correlation (-0.97) between the integration of the past history of power consumption (with $\bar{P}$ as the reference value), i.e., $\int_0^t (Nf_p(u(\tau)) - \bar{P})d\tau$ and E_QoSD(t), which is shown in Fig.1. Hence, we propose to model E_QoSD(t) with linear regression as follows:

$$\text{E\_QoSD}(t) = \alpha \int_0^t (Nf_p(u(\tau)) - \bar{P})d\tau + g(\mu_W) + \omega_1, \tag{4}$$

where $g(\mu_W)$ is the proper function that transforms the mean of queue length to mean of system degradation, α can be determined from simulation, and ω_1 is a zero mean random variable with known variance. Since $f_p(u(t)) = ku(t) + P_{idle}$ from Eq.(1), and $\bar{P} = Nf_p(\bar{u}) = N(k\bar{u} + P_{idle})$, Eq.(4) can be simplified as:

$$\text{E\_QoSD}(t) = \alpha Nk \int_0^t (u(\tau) - \bar{u})d\tau + g(\mu_W) + \omega_1, \tag{5}$$

in which we linearly transform the integration of power consumption to the integration of the processing rate.

The variance of the QoS degradation, V_QoSD(t), is affected by the number of job departures Dep(t) in every

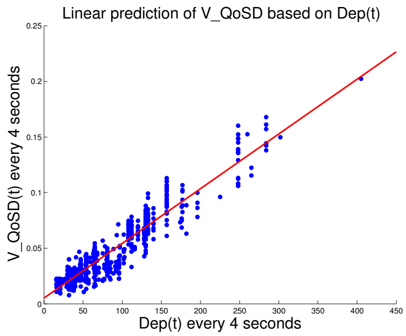

Fig. 2. Variance of QoS degradation is characterized by the number of job departure (i.e., finished) in 4 seconds. The higher departure number results in higher uncertainty to the system, and thus higher variance.

4 seconds from the observation. A larger $\text{Dep}(t)$ results in more sample uncertainties, and therefore larger $\text{V\_QoSD}(t)$. Fig.2 is the scattered plot between $\text{Dep}(t)$ and $\text{V\_QoSD}(t)$, whose correlation is 0.94. With linear regression we have:

$$\text{V\_QoSD}(t) = \beta(\text{Dep}(t) - 1) + \omega_2, \tag{6}$$

where $\text{Dep}(t)$ can be estimated as a Poisson random variable with $\lambda = Nu(t)\Delta_t$ ($\Delta_t = 4$ seconds) based on simulation results. β and ω_2 can be determined from simulation.

Given $\text{E\_QoSD}(t)$ and $\text{V\_QoSD}(t)$, the PDF of the uniformly distributed $\text{QoSD}(t)$ is

$$p(\text{QoSD}(t)) = \begin{cases} \frac{1}{\sqrt{12\text{V\_QoSD}(t)}} & \text{QoSD}(t) \in [a,b] \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

where the lower and upper bounds are

$$\begin{aligned} a &= \text{E\_QoSD}(t) - \sqrt{3\text{V\_QoSD}(t)} \\ b &= \text{E\_QoSD}(t) + \sqrt{3\text{V\_QoSD}(t)}. \end{aligned} \tag{8}$$

If the DCO signs a contract with users in which a penalty $C(\text{QoSD}(t))$ is added when the QoS degradation exceeds a pre-defined level Q, then the expected period cost per job departure incurred by $\text{QoSD}(t)$ is

$$\int_{Q}^{\infty} p(\text{QoSD}(t))C(\text{QoSD}(t))d\text{QoSD}(t) - \Pi^{SV}, \tag{9}$$

where Π^{SV} represents the credit earned from per job departure. The overall period cost of QoS degradation for every 4 seconds equals to the expected cost of all job departures in that 4 seconds:

$$PC_{QoSD}(t) = \mathbf{E}\Big[\text{Dep}(t)\Big(\int_{Q}^{\infty} p(\text{QoSD}(t))C(\text{QoSD}(t))d\text{QoSD}(t) - \Pi^{SV}\Big)\Big]. \tag{10}$$

Finally, based on (2) and (10), the total period cost function for every 4 seconds is

$$\begin{aligned} &PC_{total}(t) = PC_{track}(t) + PC_{QoSD}(t) = \\ &\Pi^{Err}|Nf_p(u(t)) - (\bar{P} + y(t)R)| + \\ &\mathbf{E}\Big[\text{Dep}(t)\Big(\int_{Q}^{\infty} p(\text{QoSD}(t))C(\text{QoSD}(t))d\text{QoSD}(t) - \Pi^{SV}\Big)\Big]. \end{aligned} \tag{11}$$

Next, we formulate the state dynamics of the DP. Clearly $\text{E\_QoSD}(t)$ in Eq.(5) is not Markov with respect to $u(t)$. We transform the variable into a memoryless one by adding an auxiliary variable $z(t)$ representing the integration of the processing rate $u(t)$ corresponding to $\bar{u}$ up to time t. Letting $z(0) = 0$, we have the dynamic of $z(t)$ as

$$z(t+1) = z(t) + (u(t) - \bar{u}). \tag{12}$$

Substituting $z(t)$ into (5), we have

$$\text{E\_QoSD}(t) = \alpha Nkz(t) + g(\mu_W) + \omega_1. \tag{13}$$

The dynamics of the RSR signal can be formulated by a Markov chain with two states: the value of $y(t)$ and the sign of $y(t) - y(t-1)$, namely $D(t)$, representing the direction of the signal changes at time t. Conceptually this can be represented as:

$$\begin{aligned} y(t+1) &= f_1(y(t), D(t)) \\ D(t+1) &= f_2(y(t), D(t)) \end{aligned}. \tag{14}$$

Since the statistic behavior of the RSR signal is known ahead, function f_1 and f_2 can be calculated in advance. They can also be mined from historical ISO RSR signal data. Detailed discussion on Eq.(14) is referred to [15].

We formulate the stochastic DP problem as a discounted cost infinite horizon DP. If we denote the value function as $J(y,D,z)$ and the overall state dynamics by

$$x(t+1) = f(x(t)), \tag{15}$$

where $x(t)$ is composed of $\{y(t), D(t), z(t)\}$, then the Bellman Equation is

$$J(y,D,z) = PC_{total}(y,D,z) + \eta\mathbf{E}\Big[J(f(y,D,z))\Big], \tag{16}$$

where η is the discounted rate.

To conclude, the state variables are $\{y(t), D(t), z(t)\}$, the control variable is $u(t)$, the disturbances are $\text{E\_QoSD}(t), \text{V\_QoSD}(t)$, and the discounted cost infinite horizon DP is to solve the following problem

$$\begin{aligned} &\min \quad (16) \text{ over } u \\ &\text{s.t.} \quad (6), (7), (12), (13), (14). \end{aligned} \tag{17}$$

III. OPTIMAL FEEDBACK CONTROL

In this section, we introduce the simulation method to solve for the optimal feedback policy of our discounted rate infinite horizon DP problem. We then discuss the policy and compare it with a heuristic operating solution in literature.

A. Simulation Method

To figure out the optimal control policy u, we solve the discounted cost infinite horizon problem with the value iteration method. We normalize and discretize the control variable $u(t)$ into 11 levels in the range of [0,1], with the granularity at 0.1. For the state variable $z(t)$ introduced in Eq.(12), we simulate extensively and study its distribution to acquire its possible range. Based on the distribution, a range of [-20, 20] can include more than 95% values of $z(t)$. For $|z(t)| > 20$, we truncate them to 20. Since $z(t)$ is the integral of $u(t)$, we discretize $z(t)$ using the same granularity as $u(t)$. We discretize the state variable $y(t)$ at

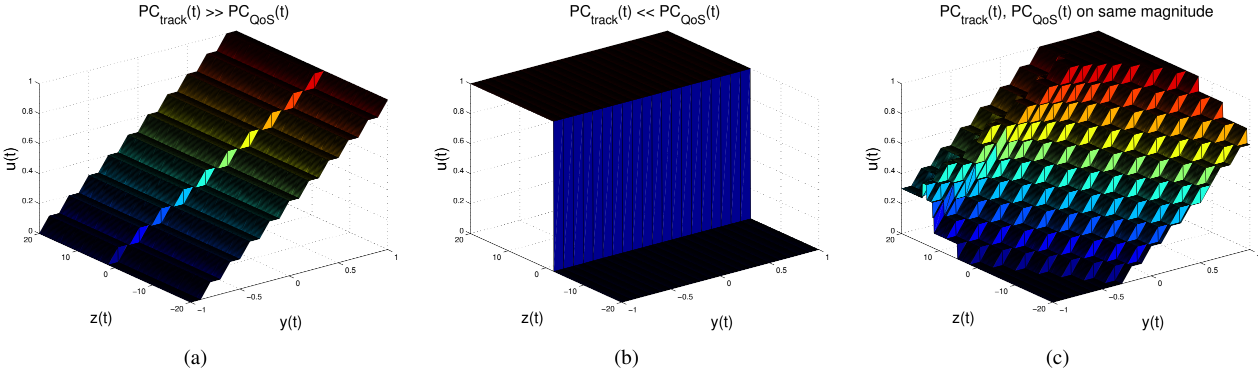

Fig. 3. The optimal policies $u(t)$ via $y(t)$ and $z(t)$ given $D(t)=1$, of three cases: (a) $PC_{track}(t) >> PC_{QoSD}(t)$, i.e., the tracking cost dominates in the overall period cost function; (b) $PC_{track}(t) << PC_{QoSD}(t)$, i.e., the cost of QoS degradation dominates in the overall period cost function; (c) $PC_{track}(t)$ and $PC_{QoSD}(t)$ are on the same order of magnitude.

the granularity of 0.1. Since the state variable $D(t) = \pm 1$, there are $16842 (= 401 \times 21 \times 2)$ different states in total.

By understanding the real-life data center service level agreement (SLA), we define the cost function of the QoS degradation in Eq.(9) as

$$C(\text{QoSD}(t)) = \begin{cases} 0, & \text{if } \ \text{QoSD}(t) \in [1, Q) \\ \Pi^D \text{QoSD}(t), & \text{otherwise} \end{cases}, \quad (18)$$

which is a discontinuous function. Π^D is a constant, representing the penalty price on per unit of the QoS degradation. We select the threshold $Q = 3$ in simulation. Π^D here and Π^{SV} in Eq.(10) are estimated based on the price information of Amazon Web Service (AWS) [20]. In general, Π^D and Π^{SV} have the same order of magnitude.

B. Optimal Policy

The optimal policy can be quite different while selecting different values of Π^{ERR} and Π^D. A large Π^{ERR} can lead to $PC_{track}(t) >> PC_{QoSD}(t)$, while large Π^D can have $PC_{track}(t) << PC_{QoSD}(t)$. Fig. 3 shows the optimal control policy u via key state space variables $y(t)$ and $z(t)$ given $D(t) = 1$, of the following three cases:

(i) Π^{ERR} is large, i.e., $PC_{track}(t) >> PC_{QoSD}(t)$. In this case the cost of tracking error is much larger than that of the QoS degradation, the optimal policy tends to always track the RSR signal $y(t)$ as accurate as possible to minimize the overall costs. So the policy is sensitive to and monotonically varies with $y(t)$, and is almost independent of $z(t)$;

(ii) Π^D is large, i.e., $PC_{track}(t) << PC_{QoSD}(t)$. In this case the cost of tracking error is much smaller than that of the QoS degradation, the optimal feedback policy is a bang-bang controller that either sets $u(t)$ at the minimal or at the maximal level. Specifically, if the mean of QoS degradation E_QoSD(t) is large because of a small $z(t)$, then $u(t) = 0$ and the policy decreases the number of departure jobs at t, i.e., Dep(t). If E_QoSD(t) is small because of a large $z(t)$, then $u(t) = 1$ and the policy increases the number of departure jobs at t, so as to minimize the overall costs. Overall, the optimal policy in this case is only sensitive to $z(t)$ and is independent of $y(t)$;

(iii) $PC_{track}(t)$ and $PC_{QoSD}(t)$ have the same order of magnitude. While **(i)** and **(ii)** are two extreme cases, case **(iii)** requires to balance between the signal tracking costs and the QoS degradation costs. From Fig. 3(c) we find that the optimal policy depends on both $z(t)$ and $y(t)$. The policy shows that: 1) for the same $z(t)$, the optimal processing rate $u(t)$ increases when the signal $y(t)$ increases, so as to better track the signal; 2) for the same $y(t)$, the optimal processing rate generally increases as $z(t)$ increases, which shows that when the mean of QoS degradation is small, the optimal policy tries to finish and depart more jobs during that moment; 3) there is a non-monotonic behavior of the optimal policy around $z(t) = 10$ to $z(t) = 15$. This region of $z(t)$ corresponds to the region of E_QoSD(t) around threshold Q in Eq.(18), which is the discontinuous turning point of the QoS degradation penalty cost function, while below which there is no degradation penalty and above which the penalty linearly increases. Such non-monotonic behavior of $u(t)$ can be explained as follows: When $z(t)$ corresponds to E_QoSD(t) that is near the left extreme of threshold Q, at which there is still no penalty of QoS degradation, the policy applies larger processing rate $u(t)$ to finish and depart more jobs to minimize the overall costs. When $z(t)$ is larger, e.g., $z(t) = 20$, however, the DCO does not necessarily use large processing rate, as there would be also no penalty if jobs are finished and depart later when $z(t+\Delta) = 15$ for small Δ. Instead, the system can focus more on eliminating tracking errors at that moment to minimize the overall costs. This explains the phenomenon that the optimal processing rate $u(t)$ can be larger for $z(t) = 15$ than $z(t) = 20$.

The optimal policy of $D(t) = -1$ is similar to that of $D(t) = 1$, except that there is a small shift in the figure along the direction of $z(t)$ axis. This is because that when $D(t) = -1$, there is a higher probability that the RSR signal $y(t)$ is going to decrease in the future than when $D(t) = 1$. In order to eliminate the overall tracking error, the optimal policy of $D(t) = -1$ prefers lower $u(t)$ than that of $D(t) = 1$. Overall, the shift is small, which shows that the policy is not very sensitive to the state variable $D(t)$.

C. Policy Comparison

We compare the optimal policy with a previous heuristic operating policy [10], which simply tracks the signal $y(t)$

as accurate as possible and does not include costs of QoS degradation in the period cost function. We consider the same scenario for both policies: a data center with $N = 1000$ servers, jobs arrive following the Poisson distribution with the arrival rate as $\lambda = 50\% N u_{max}$. We run simulation with a real 24-hour historical RSR signal data from PJM [8] as our $y(t)$, and use only data from the 2^{nd} to the 23^{rd} hour (data from the 1^{st} and the last hour is not clean and stable due to the effects of the initialization and termination of the experiment). We treat this 22-hour simulation as 22 repetitions of the 1-hour experiment and then measure the statistics in order to achieve statistical confidence. The price information used in simulation is estimated based on real PJM [8] and AWS [20] data, i.e., Π^{Err} = 0.2\$/kWh, Π^{D} = Π^{SV} = 0.1\$/h. These price values lead $PC_{track}(t)$ and $PC_{QoSD}(t)$ to share the same order of magnitude, and thus the scenario falls in the category (iii) introduced in Section III-B.

We measure the tracking cost J_{track} and the cost of QoS degradation J_{QoSD}, and recall that the overall cost J_{total} equals to $J_{track} + J_{QoSD}$. Comparing to the previous heuristic operating policy that best tracks the signal $y(t)$ [10], our DP optimal policy incurs larger J_{track}, which is expected, as the previous policy tracks signal the best. However, our new policy leads to much smaller J_{QoSD}, due to the fact that the QoS degradation is carefully taken into account in our DP solution. Overall, the total cost, J_{total} of the DP optimal policy is decreased by 4.55% on average.

IV. HOUR-AHEAD BIDDING MECHANISM

Acquiring the optimal policy, we then study the optimal hour-ahead bidding strategies for the DCO in the energy and reserve market. For the scenario of the data center with 1000 servers and the utilization of 50%, Eq.(3) provides the constraints of the bidding values. Obeying the constraints, we run simulations with different ($\bar{P}$,R) and measure the data center's hourly overall bill as

$$B(\bar{P},R) = \Pi^{E}\bar{P} - \Pi^{R}R + J_{total}(\bar{P},R), \tag{19}$$

where $J_{total}(\bar{P},R)$ is the summed value of the tracking error cost and the cost of QoS degradation in DP formulation, introduced in Section III-C.

For simplicity of notation, we denote $N f_p(0.5u_{\max})$, i.e., the lower bound of $\bar{P}$ in Eq.(3) as P_{lb}. Table I shows the overall hourly bill (in dollars) of a 1000-server data center with $\bar{P}$ = 1.001, 1.1 and 1.2 P_{lb} respectively in each row[3]. For each selected value of $\bar{P}$, the maximal possible reserve value is: $R_{\max} = \min(N f_p(u_{\max}) - \bar{P}, \bar{P} - N P_{idle})$ from Eq.(3). In the table, we measure the bill via R = 20%, 40%, 60%, 80% and 100% $R_{\max}$ respectively, for each $\bar{P}$. $\Pi^{E} = \Pi^{R}$ = 0.2 \$/kWh based on the real market data ([8], [21]).

The table shows that, satisfying the constraints in Eq.(3), the overall hourly bill of the data center in RSR provision increases monotonously as $\bar{P}$ increases, and decreases monotonously as R increases. Therefore, in order to minimize

[3]1.001P_{lb} is selected because P_{lb} itself does not satisfy Eq.(3), however, a power value sightly larger than P_{lb} is able to, e.g., 1.001P_{lb}.

TABLE I
THE HOURLY BILL OF A 1000-SERVER DATA CENTER WITH RSR PROVISION VIA DIFFERENT ($\bar{P}$, R)

	100%	80%	60%	40%	20%
1.001	\$31.33	\$34.89	\$38.23	\$41.36	\$43.90
1.1	\$40.79	\$43.55	\$46.43	\$49.58	\$53.00
1.2	\$47.99	\$49.49	\$50.82	\$52.02	\$53.19

[a]1.001, 1.1 and 1.2 represent $\bar{P}$ = 1.001P_{lb}, 1.1P_{lb} and 1.2P_{lb} respectively, with $P_{lb} = N f_p(0.5u_{\max})$.
[b]100%, 80%, 60%, 40% and 20% represent R = 100%R_{max}, 80%R_{max}, 60%R_{max}, 40%R_{max} and 20%R_{max} respectively, with $R_{max} = \min(N f_p(u_{\max}) - \bar{P}, \bar{P} - N P_{idle})$.

the monetary costs, the optimal bidding mechanism for the data center RSR provision is to choose the smallest $\bar{P}$ and the largest R that satisfy Eq.(3).

V. RELATED WORK

As the overall need for reserves increases with increasing renewable penetration, relying on conventional generators for reserves is costly and environmentally undesirable. Results show that loads, by acting as both positive and negative generation sources, can promise to respond to RSR signals to maintain grid balance and help reduce the need for secondary reserves from conventional fossil fuel generators ([22], [23]). An off-only switched control has been proposed to achieve bi-directional electricity modulation [24]. The operator at each time only needs to decide the fraction of appliances to be disconnected. Disconnected appliance will automatically reconnect after a fixed amount of time. Some recent work proposes joint optimization frameworks to minimize the average cost of RSR deployment requirement violation and consumer dis-utility, by employing dynamic system models or Markov decision process [25]. In addition to thermostatic loads, several heuristic scheduling policies have been proposed to solve for large scale participation of deferrable load in reserve provision. The proposed policies include earliest deadline first strategy, least laxity first strategy, and receding horizon control [26]. Moreover, deadline constrained appliances, such as electric vehicles, washing machines, can be coordinated to fill in the overnight demand valley, to mitigate renewable energy intermittency and reduce transmission congestion ([27], [28], [29], [30], [31], [32], [33]).

Discovering that the power consumption of data centers worldwide keeps increasing rapidly during the past few decades, a growing number of studies begin to model and investigate the capacity and benefit of enabling data centers to participate in multiple types of power market demand response and to provide various reserves, including RSR ([21], [34], [35], [36]). Various data center power management techniques, e.g., load shedding and load shifting, are exploited for data center demand response participation ([37], [38]). Brocanelli et al. [39] propose to provide reserves by exploiting synergies of degrees of freedom in data center and employee PHEVs. Finally, other work explores the opportunities of data centers in stabilizing the power network and balancing the electric power load ([40], [41], [42]).

VI. CONCLUSION

In the work, we have modeled QoS degradation of job processing in data centers as a dynamic uniform probability distribution whose dynamic mean and variance depend on recent control trajectory statistics and job departures. The mean and variance of this probability distribution serve as sufficient statistics that describe the useful information on the current QoS performance. We have introduced a stochastic DP that employs these sufficient statistics along with other state information to determine optimal state feedback policies, which enable the DCO to (i) reduce by around 5% energy cost compared to heuristic policies proposed in literature, and (ii) buy energy and sell reserve optimally in the hour-ahead market. Future work will focus on investigating additional control actions by accounting for multiple server operating modes, such as putting servers to sleep, and pursuing the characterization of optimal switching control policies, i.e., binary level control, in addition to the currently deployed continuous power consumption rate control.

REFERENCES

[1] The 2020 climate and energy package. *European Commission: http://ec.europa.eu/clima/policies/package*.

[2] *American Wind Energy Association: http://www.awea.org/Advocacy*.

[3] Annual energy outlook 2014. *U.S. Energy Information Administration: http://www.eia.gov/forecasts/aeo*.

[4] L. Rao, X. Liu, M. D. Ilic, and J. Liu. Distributed coordination of internet data centers under multiregional electricity markets. *Proceedings of the IEEE*, 100(1):269–282, 2012.

[5] J. Koomey. Growth in data center electricity use 2005 to 2010. *Oakland, CA: Analytics Press*, August, 1, 2010.

[6] J. Li and J. F. Martinez. Dynamic power-performance adaptation of parallel computation on chip multiprocessors. In *High-Performance Computer Architecture (HPCA), the 12th International Symposium on*, pages 77–87. IEEE, 2006.

[7] H. Chen, C. Hankendi, M. C. Caramanis, and A. K. Coskun. Dynamic server power capping for enabling data center participation in power markets. In *Intl. Conf. on Computer-Aided Design (ICCAD)*, 2013.

[8] PJM (2013). market-based regulation [online]. *http://pjm.com/markets-and-operations/ancillary-services*.

[9] Manual 2: Ancillary services manual, v3.26. *NYISO*, 2013.

[10] H. Chen, A. K. Coskun, and M. C. Caramanis. Real-time power control of data centers for providing regulation service. In *IEEE 52nd Conference on Decision and Control (CDC)*, pages 4314–4321, Dec 2013.

[11] M. C. Caramanis, I. C. Paschalidis, C. G. Cassandras, E. Bilgin, and E. Ntakou. Provision of regulation service reserves by flexible distributed loads. In *IEEE 51th CDC*, pages 3694–3700, 2012.

[12] I. C. Paschalidis, B. Li, and M. C. Caramanis. A market-based mechanism for providing demand-side regulation service reserves. In *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 21–26, Dec. 2011.

[13] F. P. Kelly, A. K. Maulloo, and D. K. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research society*, 49(3):237–252, 1998.

[14] I. C. Paschalidis and J. N. Tsitsiklis. Congestion-dependent pricing of network services. *IEEE/ACM Trans. Networking*, 8(2):171–184, April 2000.

[15] B. Zhang, M. C. Caramanis, and J. Baillieul. Optimal price-controlled demand response with explicit modeling of consumer preference dynamics. In *IEEE 53rd CDC*, December 2014.

[16] A. L. Ott. Experience with PJM market operation, system design, and implementation. *Power Systems, IEEE Trans. on*, 18(2):528–534, 2003.

[17] *PJM Manual 12: Balancing Operations, www.pjm.com*, 2012.

[18] PJM. *White Paper on Integrating Demand and Response into the PJM Ancillary Service Markets*, 2005.

[19] H. Chen, M. C. Caramanis, and A. K. Coskun. The data center as a grid load stabilizer. In *Design Automation Conference, 19th Asia and South Pacific (ASP-DAC)*, pages 105–112, 2014.

[20] *Amazon Web Service Pricing: http://aws.amazon.com/ec2/pricing/*.

[21] D. Aikema, R. Simmonds, and H. Zareipour. Data centres in the ancillary services market. In *International Green Computing Conference (IGCC)*, pages 1–10. IEEE, 2012.

[22] S. Meyn, P. Barooah, A. Busic, and J. Ehren. Ancillary service to the grid from deferrable loads: The case for intelligent pool pumps in florida. In *IEEE 52nd CDC*, pages 6946–6953, Dec 2013.

[23] B. Zhang and J. Bailliuel. A two level feedback system design to regulation service provision. arXiv:1301.6386, 2013.

[24] T. M. Keep, F. E. Sifuentes, D. M. Auslander, and D. S. Callaway. Using load switches to control aggregated electricity demand for load following and regulation. In *Power and Energy Society General Meeting*, pages 1–7, July 2011.

[25] D. S. Callaway and I. A. Hiskens. Achieving controllability of electric loads. *Proceedings of the IEEE*, 99(1):184–199, 2011.

[26] A. Subramanian, M. Garcia, A. Dominguez-Garcia, D. Callaway, K. Poolla, and P. Varaiya. Real-time scheduling of deferrable electric loads. In *American Control Conference (ACC)*. IEEE, 2012.

[27] J. A. P. Lopes, F. J. Soares, and P. M. R. Almeida. Integration of electric vehicles in the electric power system. *Proceedings of the IEEE*, 99(1):168–183, 2011.

[28] L. Gan, U. Topcu, and S. Low. Optimal decentralized protocol for electric vehicle charging. In *50th CDC-ECC*, pages 5798–5804. IEEE, 2011.

[29] Z. Ma, D. S. Callaway, and I. A. Hiskens. Decentralized charging control of large populations of plug-in electric vehicles. *Control Systems Technology, IEEE Transactions on*, 21(1):67–78, 2013.

[30] M. Caramanis and J. M. Foster. Management of electric vehicle charging to mitigate renewable generation intermittency and distribution network congestion. In *IEEE 48th Conference on Decision and Control, jointly with the 28th Chinese Control Conference (CDC/CCC)*, pages 4717–4722, Dec. 2009.

[31] W. Kempton and J. Tomić. Vehicle-to-grid power implementation: From stabilizing the grid to supporting large-scale renewable energy. *Journal of Power Sources*, 144(1):280–294, 2005.

[32] Y. Gu and L. Xie. Look-ahead coordination of wind energy and electric vehicles: A market-based approach. In *North American Power Symposium (NAPS)*, pages 1–8. IEEE, 2010.

[33] N. Rotering and M. Ilic. Optimal charge control of plug-in hybrid electric vehicles in deregulated electricity markets. *Power Systems, IEEE Transactions on*, 26(3):1021–1029, 2011.

[34] H. Chen, M. C. Caramanis, and A. K. Coskun. Reducing the data center electricity costs through participation in smart grid programs. In *IGCC*, pages 1–10. IEEE, 2014.

[35] M. Ghamkhari and H. Mohsenian-Rad. Data centers to offer ancillary services. In *Smart Grid Communications (SmartGridComm), 3rd Intl. Conf. on*, pages 436–441. IEEE, 2012.

[36] M. Ghasemi-Gol, Y. Wang, and M. Pedram. An optimization framework for data centers to minimize electric bill under day-ahead dynamic energy prices while providing regulation services. In *IGCC*, pages 1–10. IEEE, 2014.

[37] A. Wierman, Z. Liu, I. Liu, and H. Mohsenian-Rad. Opportunities and challenges for data center demand response. In *IGCC*, pages 1–10. IEEE, 2014.

[38] G. Ghatikar, V. Ganti, N. Matson, and M. A. Piette. Demand response opportunities and enabling technologies for data centers: Findings from field studies. *LBNL-5763E. pdf*, 2012.

[39] M. Brocanelli, S. Li, X. Wang, and W. Zhang. Joint management of data centers and electric vehicles for maximized regulation profits. In *IGCC*, pages 1–10. IEEE, 2013.

[40] Z. Liu, I. Liu, S. Low, and A. Wierman. Pricing data center demand response. In *ACM SIGMETRICS*, 2014.

[41] H. Wang, J. Huang, X. Lin, and H. Mohsenian-Rad. Exploring smart grid and data center interactions for electric power load balancing. *ACM SIGMETRICS Performance Evaluation Review*, 41(3):89–94, 2014.

[42] R. Wang, N. Kandasamy, C. Nwankpa, and D. R. Kaeli. Data centers as controllable load resources in the electricity market. In *Intl. Conf. on Distributed Computing Systems*, 2013.