

Providing Regulation Services and Managing Data Center Peak Power Budgets

Baris Aksanli and Tajana Rosing
Computer Science and Engineering Department
University of California, San Diego
{baksanli, tajana}@ucsd.edu

Abstract— Data centers are good candidates for providing regulation services in the power markets due to their large power consumption and flexibility. In this paper, we develop a framework that explores the feasibility of data center participation in these markets. We use a battery-based design that can not only help with providing ancillary services, but can also limit peak power costs without any workload performance degradation. The results of our study using data for a 21MW data center show up to \$480,000/year savings can be obtained, corresponding to 1280 more servers providing services.

I. INTRODUCTION AND RELATED WORK

Substantial integration of electric vehicles and renewable energy sources into the electric grid poses significant challenges to ensuring the grid stability. Ancillary services are one mechanism utility companies use to ensure stability. It includes multiple mechanisms, such as demand-response (DR), spinning and non-spinning reserves and regulation services. DR is used to motivate a reduction in consumer demand with price incentives. Spinning reserves are similar to DR but include explicit contracts with service providers. This paper focuses on contract-based regulation services, which are used to balance the demand and supply. For example, when there is high energy generation, the utility might need higher power demand from consumers. Similarly, it might request users to reduce their consumption when the power generation is at a premium.

Recently, utilities have allowed non-generator sources to participate in regulation markets [1]. These sources should have large consumption and some power flexibility to allow adjustments. Power consumption of data centers is growing rapidly, up to 100MW per individual site [2]. The data center's ability to adjust its power demand at run time by employing techniques such as dynamic voltage-frequency scaling (DVFS), virtual machine (VM) migration and peak power shaving make them a good choice for regulation services. While both DVFS and VM migration have some performance overhead, recent battery-based peak shaving techniques [3] [4] [5] are capable of reducing the power consumption at *no cost to performance*. Significant savings, of up to \$75K/month for a 10MW data center, can be obtained with battery-based peak shaving. However, none of these publications consider the feasibility of using the energy storage in data centers to participate in the regulation markets. This is one key contribution of our work.

There are a few recent studies that investigate the data center participation in the ancillary services market [6] [7] [8] [9]. Wang et al. [9] model a distributed set of data centers and explore DR using VM migration. Ghamkhari et al. [8] analyze the savings of a single data center participating in voluntary load reduction. Chen et al. [7] explore the ability of data centers to provide regulation services using dynamic server power capping techniques. Aikema et al. [6] analyze a number of different ancillary services for data centers, including DR, spinning reserves and regulation.

All previous studies interfere with the workload behavior resulting in performance degradation, which is a big issue for response-time critical workloads. They also do not consider peak power costs. When providing regulation services, the utility is given the prerogative to demand a change in the power consumption of the data center by as much as a maximum amount specified in the contract over the given interval. This amount, if not properly handled, may raise the peak power level of the data center, and increase the utility bill. The data center should adjust its average power demand and the regulation capacity to be allocated to ensure that the peak power costs do not eliminate the savings from providing regulation services.

We propose a framework that analyzes the data center participation in the regulation markets while also considering the peak power budgets. We use a battery-based peak shaving design to avoid performance penalty to workloads. We study two common battery types, lead-acid (LA) and lithium iron phosphate (LFP). Our framework consists of two cases with different peak power assumptions. We present a method for each case, which first analyzes if providing regulation is feasible and if so, shows how the regulation capacity should be allocated to maximize savings. Normal peak shaving takes place to limit peak power costs if the data center chooses not to participate. We leverage the data from NYISO and CAISO markets to show the effectiveness of our framework. Our results show that for a 21MW data center, up to \$480,000/year savings can be obtained using our methods, equivalent to 1280 more servers operating, and 5.08% increase in data center profit percentage. Also, if peak power costs are not considered with regulation, the savings can be miscomputed by up to 385%.

II. BACKGROUND

Providing regulation services requires the agreement of both the regulation service provider, in our case data centers, and the independent system operator (ISO), which provides the electricity to end-users. This service can take place in real-time, hour-ahead or day-ahead markets, which all have different pricing schemes and timing requirements. In a given service interval, the regulation service provider should determine its average power demand, P_{ave} , and the regulation capacity, C_{reg} , it can provision in that interval. It then gives this information to the utility, agreeing that the utility can issue fine-grained signals that can change the power consumption of the service provider to any value within the interval $[P_{ave} - C_{reg}, P_{ave} + C_{reg}]$ uniformly. For data centers, P_{ave} depends on the resource utilization and it is typically around 50% [10] and C_{reg} changes based on the load flexibility [7]. Some previous studies, e.g. [6], incorrectly assumed that the power demand of a data center providing regulation service does not change within a service interval. When the power is set to a value in $[P_{ave} - C_{reg}, P_{ave}]$, a power shaving method can be used. Utility may also demand the power be increased from P_{ave} to a value in $[P_{ave}, P_{ave} +$

c_{reg}] to balance its larger energy supply and smaller demand. This may lead to a conflict between a need to keep peak power under a threshold due to electricity pricing [3] vs. the need to respond to utility controls as a part of the regulation contract.

In this work, we assume that the data center already uses a battery based peak power shaving method as described in [5], which avoids the performance degradation of traditional power shaving methods. Also, this method has longer battery lifetime and smaller communication overhead than the other designs. The latter is needed to ensure that the data center receives the best regulation service rates [7]. In contrast to previous studies, we do not interfere with any jobs, but instead control the battery output to track the actual and the targeted power demand.

In battery based peak power shaving, the data center first determines a peak power threshold, P_{th} and then discharges the batteries when the actual demand is over P_{th} and recharges the batteries during lower demand. The physical properties of the batteries influence the choice of this threshold. Previous studies [3] [5] show that if the batteries are discharged deeply or with high discharging currents, their expected lifetime decreases. Hence, first a fixed limit for battery depth-of-discharge (DoD) is found that is economically feasible, then the peak power limit is estimated based on this DoD limit. Typical DoD limit values are 20-40% for LA batteries and 60% for LFP batteries [3].

III. DATA CENTERS PROVIDING REGULATION SERVICES

Data centers that leverage batteries to shave peak power can respond to utility commands for regulation by changing the battery charge and discharge intervals, thus requiring only minor changes to the already implemented battery control system. We analyze two different types of data center operation where both peak shaving and regulation controls are present. The first case does not alter the average data center power demand as in the previous studies [6] [7], but increases the peak power threshold to match the allotted regulation capacity. Different than the previous studies, we show how the decision mechanism should be designed to consider peak power costs. Our other solution is different than previous studies. It does not change the power threshold but create flexibility in data center power consumption by adjusting the access to the batteries. This is the key point not to degrade workload performance.

Fixed Average Power: Careful control of data center batteries can ensure that the data center can keep its average power equal to the peak power threshold, $P_{ave} = P_{th}$. If the regulation capacity is C_{reg} , then the data center power consumption can be any value in $[P_{ave} - C_{reg}, P_{ave} + C_{reg}]$ so the peak power becomes $P_{th} + C_{reg}$, instead of P_{th} . The savings from regulation need to be larger than the difference between the original and the elevated peak power cost. The peak power cost is charged with the largest consumption over a month, and thus, it increases with the maximum regulation capacity. Then the condition becomes: $C_{reg} * c_r \geq (P_{th} + C_{reg_{max}}) * c_{pp} - P_{th} * c_{pp}$, where c_r is the hourly regulation price in \$/MW, $C_{reg_{max}}$ is the maximum regulation capacity over a month in MW, c_{pp} is the monthly peak power cost in \$/MW, around \$12,000/MW [4]. This constraint is defined for an hourly interval, while the peak power cost is charged on a monthly basis. We assume that the data center provides regulation in each interval over a month and average the monthly peak power costs over all the intervals to obtain a lower bound of the average regulation price:

$$c_{r_{ave}} \geq \frac{c_{pp}}{30*24} * \frac{C_{reg_{max}}}{C_{reg_{ave}}} \quad (1)$$

where $c_{r_{ave}}$ is the average monthly regulation price in \$/MW and $C_{reg_{ave}}$ is the average regulation capacity provisioned over a month in MW. Eq. 1 assumes that the regulation should be provided in each interval over a month with a price larger than $c_{r_{ave}}$ to make up for the extra peak power cost. A solution to this is to limit $k = C_{reg_{max}}/C_{reg_{ave}}$, thus limiting the $C_{reg_{max}}$ so that the possibility of reaching a very high peak power level is eliminated. Then, $C_{reg_{max}}$ is limited by $C_{reg_{ave}} * k$, preventing the peak power from exceeding a predefined limit.

After setting a lower bound on the regulation price, the next step is to set the regulation capacity, C_{reg} . The upper bound of the power flexibility interval, $P_{th} + C_{reg}$, cannot be greater than the nominal data center power demand in that interval, P_{nom} , when the batteries are discharging. Note that at this point, if data center does not provide regulation services, the batteries discharge to reduce the data center power demand from P_{nom} to P_{th} . This adjusted consumption can be increased by stopping some batteries discharging, up to P_{nom} . In addition, the lower bound, $P_{th} - C_{reg}$ cannot be smaller than P_{nom} when the batteries are charging. Then, we can determine the regulation capacity in each interval, t , as:

$$C_{reg}(t) = \begin{cases} \min(P_{nom}(t) - P_{th}, C_{reg_{max}}), & \text{if } P_{nom}(t) > P_{th} \\ \min(P_{th} - P_{nom}(t), P_{peak} - P_{th}, C_{reg_{max}}), & \text{otherwise} \end{cases} \quad (2)$$

Equation 2 ensures that the power demand from the utility stays within the acceptable range and C_{reg} does not exceed $C_{reg_{max}}$. Since the average power from the utility does not change the expected battery lifetime does not change.

Varying Average Power: We analyze this option in two parts. First, we focus on intervals with batteries discharging. Since the goal is to provide regulation without increasing the original peak power, the upper limit of the regulation interval, $P_{ave} + C_{reg}$, should not exceed the peak power threshold, P_{th} . Thus, the data center should reduce its average power demand further than P_{th} when providing regulation services. Although the peak power cost does not increase, the batteries may need to discharge deeper than the allowed DoD limit to create the power flexibility, thus decreasing the expected battery lifetime and increasing the battery costs. We use the coulomb counting model described in [5] to estimate the battery costs. It models the effect of each charge/discharge cycle on the battery lifetime, based on the DoD level and the discharging current in that cycle, and calculates the cost of each cycle.

We assume that, for a given DoD level, the best peak power threshold is achieved to minimize the peak power costs. Thus, it is not possible to reduce the average power consumption further than the original best peak power threshold without violating the DoD limit for some batteries. We need to investigate the tradeoff between increasing the DoD limit in an interval to provide regulation and regulation savings. Since we do not increase P_{th} , the average power the data center reports to the utility in interval t , $P_{ave}(t)$, should be less than P_{th} . Thus, C_{reg} can be at most $P_{th} - P_{ave}(t)$ and the savings become $(P_{th} - P_{ave}(t)) * c_r(t)$ where $c_r(t)$ is the regulation price in interval t .

Some batteries may need to discharge further than the DoD limit, D , to account for the extra power demand in interval t , $P_{th} - P_{ave}(t)$. We distribute this extra demand to all batteries to minimize the extra DoD and to limit their discharging current. We obtain the extra DoD in the interval t , D_{extra} , as:

$$\frac{P_{th} - P_{ave}(t)}{N * V * C_{eff}} * |t| * 100 \quad (3)$$

where N is the number of batteries, V is the single battery voltage and C_{eff} is the effective battery capacity calculated as in [5]. We distribute the required battery power to all batteries and discharge them with the same current each time as in [5], and thus, assume that their expected lifetime is the same. Then, we calculate the cost of discharging all batteries up to DoD value $(D + D_{extra})$, rather than D , in one cycle, as:

$$Cost_{bat_{extra}} = (C_R * c_{bat}) * \left(\frac{1}{lt(D+D_{extra}, I_{d_{extra}})} - \frac{1}{lt(D, I_d)} \right) * N \quad (4)$$

where c_{bat} is the unit battery cost in \$/Ah, C_R is the single battery capacity, $I_{d_{extra}}$ is the single battery discharging current when providing regulation and I_d is the original single battery discharging current in the interval t . In equation 4, the crucial part is the function $lt(D, I)$ which calculates the expected battery lifetime (in cycles) when the battery is used with D DoD limit and I discharging current. This function considers type-specific battery properties and penalizes higher DoD values and discharging currents to reflect their negative effects on battery lifetime. More details of $lt()$ can be found in equation 4 of [5]. In equation 4, we compute the cost of using all batteries in one cycle with $(D + D_{extra}, I_{d_{extra}})$ and (D, I_d) and take the difference. This battery cost should be smaller than the regulation savings. Section IV shows that current regulation prices cannot compensate for the elevated battery costs.

Next, we focus on the intervals where the batteries are idle or recharging. The difference between $P_{nom}(t)$ and P_{th} can provide the required power flexibility. We reschedule the battery recharge events to create this flexibility. Thus, we select $C_{reg} = (P_{th} - P_{nom}(t))/2$ and $P_{ave} = (P_{th} + P_{nom}(t))/2$. But, the data center cannot provide regulation services during all such intervals since it has to ensure that batteries have enough energy stored before being discharged. We need to determine which subset of intervals where $P_{nom}(t) < P_{th}$ should be selected to provide regulation. We shift the recharge intervals such that the intervals with high regulation price provide regulation and the others make sure that batteries are recharged. Also, a recharge interval should not be shifted beyond a discharge event since that recharge event might be necessary to prevent the peak in the relevant discharge interval. This method can obtain savings even in a price conservative market, as it does not increase peak power costs or battery costs.

IV. RESULTS

We model a large data center with 50,000 Sun Fire servers, each at 175W idle and 350W peak power [3], and use a linear, CPU-utilization power consumption function. We consider non-server power consumption with the power usage efficiency (PUE) metric and use 1.2 for this value, which corresponds to an energy efficient data center [10]. We use three different data center workloads, Google Search, Orkut [14], and Facebook MapReduce [15], and scale the workloads for 50,000 servers. The breakdown of these workloads for seven days is shown in Figure 1. Then, we use an event-based simulator with these workloads to extract the total data center power demand profile.

We use battery-based peak shaving and assume a battery for each server, with 40Ah capacity, as in [5]. The battery types we use in our work are LA and LFP, from [4] [3] [5], with unit cost \$2/Ah and \$5/Ah respectively. We use multiple types of batteries since they have different optimum DoD levels [3] for peak shaving and we show how different DoD levels affect the regulation efficiency. We target the day-ahead regulation

market with higher prices than the others. Data centers can estimate the expected load for the day ahead, which can allow them to provision their resources accordingly [7]. We use pricing from NYISO and CAISO to show the importance of the market data center participates in. Figure 2 shows the NYISO numbers from [6] and CAISO ones from their database [12].

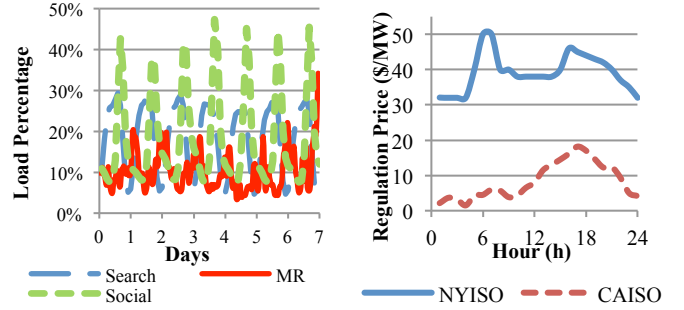


Fig. 1. Data center load composition

Fig. 2. Regulation prices

Fixed Average Power: This method does not change the average data center power demand when providing regulation, but increases the peak power. It can be seen as a representative of the previous studies [6] [7], but it differs strictly from them as it also considers peak power budgets. We first estimate the total savings for different DoD levels and k values with both NYISO and CAISO prices. Table I shows the maximum k values that can be obtained with varying DoD limits.

TABLE I. MAXIMUM k VALUES FOR DIFFERENT DoD LEVELS

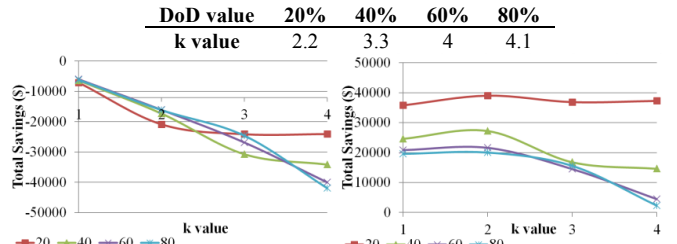


Fig. 3. Total savings result with CAISO (left) and NYISO (right) prices

Figure 3 shows the total savings in CAISO and NYISO markets. In both graphs, the x- and y-axes show changing k values and the total savings in a month in dollars. Each line corresponds to a different battery DoD level. The savings are calculated as the difference between the profit from providing regulation and the cost of increased peak power. We select the regulation capacity using different k values based on equation 2. Since the average power consumption does not change compared to the no-regulation case, there is no extra battery cost. The data center can obtain savings for any k value in the NYISO market, whereas the CAISO prices do not lead to any savings, which means that the data center should not participate in the regulation market. The best DoD value is 20% for the NYISO and 60% for the CAISO markets. The maximum savings are \$40,000 with NYISO, corresponding to 5% savings overall the electricity bill. These savings can also accommodate 1280 more servers within the same power budget.

TABLE II. ERROR PERCENTAGES WITHOUT PEAK POWER COSTS

k value	1	2	3	4
NYISO	37	52	66	79
CAISO	182	257	324	385

Table II shows the average error in savings if peak power costs are not considered with different k values. We see that the error is smaller with lower k values since k value limits the maximum regulation capacity and its effects on the peak power

cost. The error is up to 385% for the CAISO market and 80% for the NYISO. It is much higher in CAISO as peak power costs are much larger than regulation savings.

Varying Average Power: This method does not modify the peak power level but instead change the average power using the battery charge and discharge events. We first focus on the intervals with batteries discharging and assume that the original peak levels with a given battery configuration should not be increased to avoid high peak power costs. The regulation flexibility interval should be created under the original best peak power threshold. Thus, batteries discharge further than their allowed DoD limit to create this flexibility. The best peak shaving with LA and LFP batteries is obtained with 40% and 60% DoD at 17.2MW and 16.7MW respectively for our system. The expected battery lifetime is 2.5 years for LA and 6.4 years for LFP batteries.

TABLE III. REGULATION PRICE ANALYSIS FOR DISCHARGE INTERVALS

C_{reg} (MW)	D_{extra}	Bat. Life (years)		Min Reg. Price (\$/MW)	
		LA	LFP	LA	LFP
0.5	2.08	2.4	6.2	183	138
1	4.17	2.3	6	191	143
1.5	6.25	2.2	5.8	199	148

Table III shows the minimum regulation price for a given regulation capacity to compensate for the increased battery cost due to deeper discharges. We change the amount of regulation capacity and calculate the extra DoD level required. This extra DoD leads to a higher cycle cost. We compute it by estimating the battery lifetime if the battery is used with the extra DoD in each cycle. Lastly, we calculate the minimum regulation prices in \$/MW that makes up for the increased cycle cost. Table III shows that the required minimum regulation prices are much higher than the actual prices (both NYISO and CAISO), and they increase as regulation capacity rises. This method is not feasible for data centers with peak shaving and battery lifetime limitations with the current regulation prices. It can become feasible with lower battery prices or less nonlinear battery behavior. Emergency DR events, with prices up to \$500/MW [6], might be a good target to compensate for high battery costs.

Next, we investigate intervals when the batteries are idle or recharging. We use recharge shifting, shown in Section III, to obtain the intervals in which regulation can be provided. Table IV shows the monthly savings for both 40% and 60% DoD levels, in both NYISO and CAISO markets. NYISO savings are almost 3x higher than CAISO due to the higher regulation prices. Lower DoD limits lead to more than 2x savings, since the recharging intervals are more flexible. The advantages of this method are that it does not increase the original peak threshold, it does not put extra burden on the batteries, and it can obtain savings in more price-conservative markets.

TABLE IV. MONTHLY SAVINGS USING RECHARGE SHIFTING

DoD %	NYISO savings (\$)	CAISO savings (\$)
40	33628	11312
60	14600	5132

Figure 4 shows recharge shifting for NYISO (left) and CAISO (right) for a sample day. In both graphs x-axis shows the time in hours and y-axis is the power consumption in MW. The straight-line is the nominal power, the dashed line is the adjusted power using batteries without recharge shifting and dotted line shows the consumption with recharge shifting. For both graphs, recharge shifting is visible between hours 4-15 and 20-24. In 4-15hrs in the NYISO market, the data center first

decreases its consumption until $t=11.5$ to create flexibility for regulation. Then in 11.5-15hrs, it increases the consumption to complete the recharge. With CAISO prices, it stops recharging in 4-8.5hrs and then provides regulation in 8.5-15.5hrs by recharging. By recharge shifting, the data center provides regulation in the intervals with higher regulation prices. The data center has the same recharge shifting in both NYISO and CAISO in 20-24hrs due to same pricing trend in both markets.

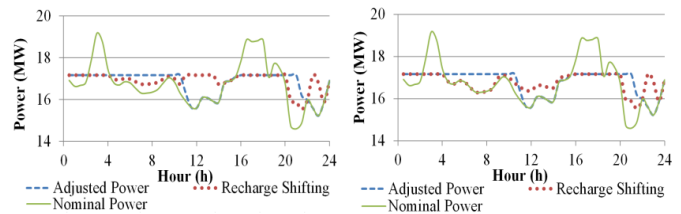


Fig. 4. Recharge shifting for NYISO (left) - CAISO (right) for a single day

V. CONCLUSION

Data centers can participate in regulation markets with their high energy consumption and ability to create flexibility in their demand to obtain savings. Previous studies all degrade the workload performance to provide regulation but do not consider peak power costs, which may result in up to 385% error in savings. Our solution adopts a battery-based peak shaving method with no performance impact on the workloads and considers the peak power costs. Increasing the peak power limits may be feasible with high regulation prices, but in a more conservative market, the best practice is to provide regulation only when the batteries recharge. We can obtain \$480,000/year savings for a 21MW data center, which can accommodate 1280 more servers and increase the profit percentage by 5.08%.

ACKNOWLEDGEMENT

This work was sponsored in part by Google, Microsoft, NSF grant 812072 and TerraSwarm, one of six centers of STARnet, a Semiconductor Research Corporation program sponsored by MARCO and DARPA.

REFERENCES

- [1] CAISO - Non-generator sources. <http://www.caiso.com/informed/Pages/StakeholderProcesses/CompletedStakeholderProcesses/NonGeneratorResourcesAncillaryServicesMarket.aspx>
- [2] G. Cook and J. Horn, "How Dirty Is Your Data? A Look at the Energy Choices that Power Cloud Computing," 2011.
- [3] V. Kontorinis, L. Zhang, B. Aksanli, J. Sampson, H. Homayoun, E. Pettis, D. Tullsen and T. Rosing, "Managing Distributed UPS Energy for Effective Power Capping in Data Centers," *ISCA*, 2012.
- [4] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, "Benefits and Limitations of Tapping into Stored Energy For Datacenters," *ISCA*, 2011.
- [5] B. Aksanli, E. Pettis, and T. Rosing, "Architecting Efficient Peak Power Shaving Using Batteries in Data Centers," *MASCOTS*, 2013.
- [6] D. Aikema, R. Simmonds, and H. Zareipour, "Data Centres in the Ancillary Services Market," *IGCC*, 2012.
- [7] H. Chen, C. Hankendi, M. C. Caramanis, and A. K. Coskun, "Dynamic Server Power Capping for Enabling Data Center Participation in Power Markets," *ICCAD*, 2013.
- [8] M. Ghamkhari and H. Mohsenian-Rad, "Data Centers to Offer Ancillary Services," *SmartGridComm*, 2012.
- [9] R. Wang, N. Kandasamy, C. Nwankpa, and D. R. Kaeli, "Datacenters as Controllable Load Resources in the Electricity Market," *ICDCS*, 2013.
- [10] L. A. Barosso and U. Holzle, *The data center as a computer: An Introduction to the Design of Warehouse-Scale Machines*, 2nd ed., 2013.
- [11] D. Meisner, C.M. Sadler, L.A. Barroso, W.D. Weber, and T.F. Wenisch, "Power management of online data-intensive services," in *ISCA*, 2011.
- [12] California ISO. <http://oasis.caiso.com/>
- [13] NYISO. www.nyiso.com
- [14] Google Transparency Report. <http://www.google.com/transparencyreport/traffic>
- [15] Y. Chen, A. Ganapathi, R. Griffith, and R. Katz, "The case for evaluating MapReduce performance using workload suites," 2011.