Multi-market Optimization of a Data Center without Storage Systems

Yangyang Fu[1] Wangda Zuo[1,2] Kyri Baker[1,2]

[1]Department of Civil, Architectural and Environmental Engineering, University of Colorado Boulder, USA,
`{yangyang.fu,wangda.zuo,kyri.baker}@colorado.edu`
[2]National Renewable Energy Laboratory, USA

Abstract

Data centers have numerous opportunities to participate in demand response programs considering their large capacities, flexible working environments and work loads, redundant design and operation, etc. Frequency regulation, as one service provided in demand response programs, can also benefits the data centers. This paper aims to develop a real-time multi-market optimization framework for a data center without storage systems to maximize their benefits from participating in both the energy market and the regulation market. Then a case study is conducted to numerically investigate the optimal bids at each hour by considering the energy cost, demand costs, and regulation revenues using a virtual data center located in PJM. Simulation results show that the proposed multi-market optimization framework can help data centers maintain minimum costs by getting maximum regulation revenues while satisfying energy and demand goals.

Keywords: Frequency Regulation, Data Center, Multi-market Optimization

1 Introduction

Data centers have numerous opportunities to participate in demand response (DR) programs considering their large energy capacities, flexible working environments and work loads, redundant design and operation, etc. For example, researches have shown that an optimized 30 MW data center is comparable to 7 MWh large-scale storage in providing DR service for the power grid (Wierman et al., 2014). Besides, some delay-tolerant data centers are allowed to have flexible work environment and workloads. What's more, the redundant design in data centers to meet reliability standards in order to guarantee their uptime and performance (Standards et al., 2005) can provide extra potentials to DR-related controls.

Frequency regulation (FR), as one type of DR, is an ancillary service that provides continuous, rapid, and automatic corrections for changes in electricity generation or use on a second-to-second basis in order to maintain the system frequency at its nominal value (e.g., 60 Hz in U.S.). Typically, FR resources are generators. FR uses certain amount of generators (e.g., about 1% of total generation) to continuously track the demand variations. The frequency must be strictly maintained within a very narrow range in order to comply with the control performance standards and the balancing authority area control error limit reliability criteria. Besides generators, fast-ramping demand side resources (DSRs) in buildings can also provide FR service to the grid by harnessing the demand flexibility provided by the modulating loads. Typical modulating loads on building side include energy storage systems such as flywheels, batteries and compressed-air energy system, electric boilers and heaters, and independent systems with variable frequency drivers (VFDs).

Recently, awareness of these potentials has drawn attention to the capabilities of data centers to participate in DR programs. A survey conducted by the Lawrence Livermore National Laboratory in 2015 shows that about 50% of the participating data centers have interest in smart pricing demand side programs, such as load shedding to avoid peak demand (Bates et al., 2015). However, data centers are reluctant to participate in fast demand response programs such as providing frequency regulation (FR) in ancillary service market, for multiple reasons. One reported concern is that data centers are still learning the process of providing FR and that providing grid services on such a fast timescale can be “outside of their visibility or control" (Bates et al., 2015). This concern is well-founded considering that these programs provide novel and relatively unexplored territory from the point of view of traditional data center control and operations.

This paper aims to explore data centers' ability of providing frequency regulation service to grids and maximize their benefits from participating regulation market and energy market as a whole. First, a synergistic control strategy together with a new regulation flexibility factor is proposed to enable the provision of regulation services in data center. Then, a real-time optimization framework is developed to maximize the data centers' benefits from participating in both the regulation market and the energy market. In Section 4, the optimization framework is evaluated in a Modelica-based environment for typical days in January and July.

$$R_{revenue} \qquad {}^{t \quad t}_{t} \quad p_{rm}\ t\ C_{reg}\ t\ dt \tag{26}$$

where p_{rm} is the real-time price signal from the regulation market, and $C_{reg}\ t$ is the regulation capacity bid for each time step.

The price signals such as p_{em} and p_{rm} need to be predicted one optimization step ahead, e.g. 1 hour in this study. Many researches have been conducted for this purpose. In this paper, historical prices of these two electrical markets are used, which means the hourly ahead prices are assumed to be perfectly predicted. The demand limit $P_{dm,lim}$ can also be predefined by the data center operators based on historical operation conditions. The maximum regulation capacity at each optimization step is set to 798.2 kW (20% of the nominal power). Note this maximum regulation capacity setting is not the feasible capacity the data center can provide at each hour, because the regulation capacity is related to data center operational conditions such as arrival rate, and weather conditions etc. This simplification has limited influence on the optimization results when the lower limit of performance score s_l is set to a high value, because if the data center makes a bid that exceeds its capacity, it cannot track the reference signal, thus the regulation performance will be low. By setting s_l to a high value can help data center make a reasonable bids when the regulation capacity is hard to predict. The optimization problem is solved using the pattern search algorithm in the optimization engine, GenOpt (Wetter et al., 2001).

4 Case Study

A data center as shown in Figure 2 is used to investigate the benefits from participating in different electrical markets. The data center is considered as a price taker only. This case study investigates the maximum benefits that data centers can obtain from both the real-time energy market and the regulation market in PJM. For the regulation service, only dynamic regulation is studied here, because its price is usually much higher than traditional regulation.

4.1 Case Description

The data center is located in Chicago, which is in ASHRAE Climate Zone 5A and within the PJM market territory. For the cooling system, there are two chillers and one integrated waterside economizer providing cooling to the data center room. This cooling system can operate in three modes: Free Cooling (FC) mode when only the WSE is enabled for cooling, Partial Mechanical Cooling (PMC) mode when the chiller and WSE are both triggered, and Full Mechanical Cooling (FMC) mode when only the chiller is activated. There are also two cooling towers, two constant-speed condenser water pumps, two variable-speed chilled water pumps, and one variable speed fan. The cooling system and its control are modelled using

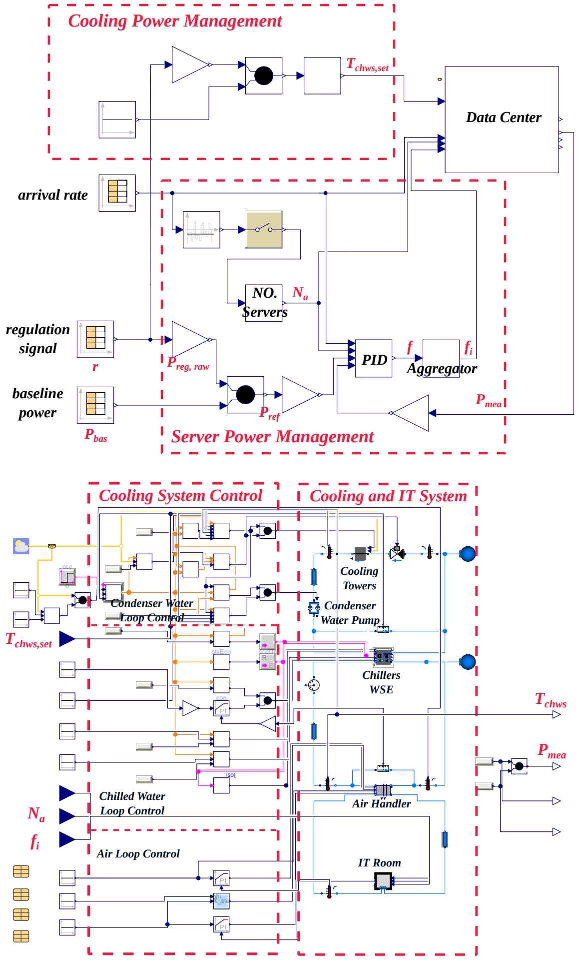

Figure 2. Modelica implementation of the studied data center for FR service: FR controller (top) and data center system (bottom)

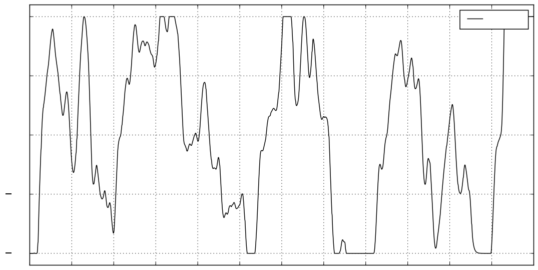

Figure 3. An example of one-hour historical RegD signal in January

an open-source equation-based Modelica environment (Fu et al., 2018, 2019a,c,b).

For the IT system, the design number of servers is 8000. The design factor γ is set to 1.5 (Li et al., 2013). The total nominal electrical load is about 2700 kW. The calibrated coefficients for Eq. (1) are $b_0 = 0.0154$, $b_1 = 1.5837$, $b_2 = 0.1373$, $c_0 = -22.3540$ and $c_1 = 121.0212$ using the method mentioned in Ref. (Li et al., 2013). When not providing FR, the server aggregator operates at a frequency of 0.8 with a regulation flexibility factor of 1.0, and the CHWST setpoint is set to 8 °C. For the internet data center, the constants C_A and C_B are set to 1 as in Ref. (Li et al., 2013).

For the multi-market optimization, all the settings are the same as the baseline except that an additional FR controller as designed in Section 2 is used to provide regulation service for the grids by adjusting the CPU frequency and CHWST setpoint. The FR flexibility factor is set to 1.1 when providing regulation services. The QoS when providing regulation services is guaranteed by constraining the average response time of the data center service to 6 ms. The lower limit of the performance score in PJM to disqualify a regulation resource is 0.4 (LLC, 2019). Here we set it to a higher value, 0.9. The real-time optimization is performed at a one-hour interval for 2 days in both January ($1/20 \sim 1/21$) (when cooling system operates at FC mode) and July ($7/20 \sim 7/21$) (when cooling system operates at FMC mode).

The price signals of the real-time energy market and the regulation service market in January and July 2018 are posted in Ref. (PJM, 2019), and the price during the optimization period is plotted as shown in Figure 4. An example of one-hour historical RegD signal is plotted in Figure 3. A real-time web service in Wikipedia (Wang et al., 2019) is used as the workload arrival profile during optimization, which is shown in Figure 5.

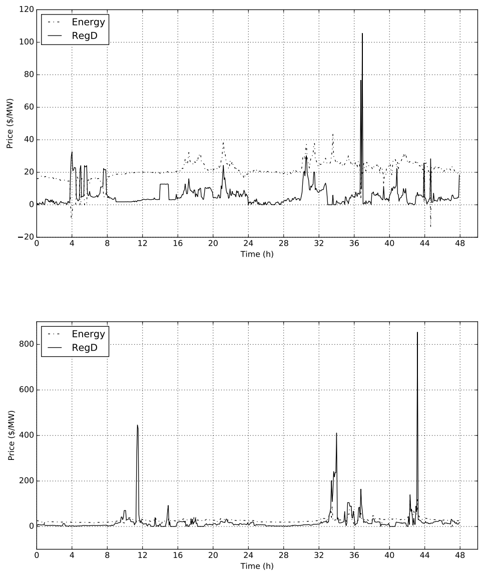

Figure 4. Historical real-time prices of PJM energy market and regulation market in January (top) and July (bottom)

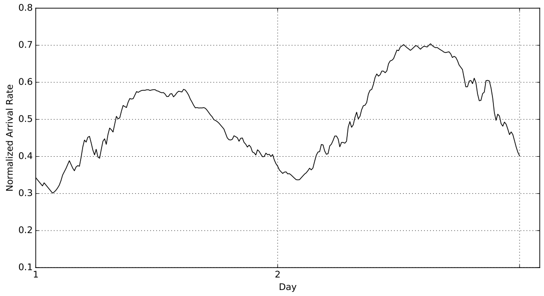

Figure 5. Two-day historical arrival rates in the data center

4.2 Results and Discussions

Table 1 compares the total cost of the data center in terms of baseline operation and multi-market optimization. The baseline system is denoted as *Base*, and the multi-market optimization is denoted as *OPT*. In both January and July, the data center without energy storage systems, using the proposed optimization framework, can benefit from participating in both energy market and regulation market. In the two days considered, *OPT* can save \$123.6 in July, while the saving is \$24.8 in January.

Table 1. Multi-market optimization of data centers

Costs	January		July	
	Base	*OPT*	*Base*	*OPT*
Energy Cost (\$)	1043.3	1042.9	1591.4	1590.6
Demand Cost (\$)	10459.3	10457.5	12063.9	12062.8
Regulation Revenue (\$)		22.6		121.7
Total Cost (\$)	11502.6	11477.8	13655.3	13531.7
Total Savings (\$)		24.8		123.6

The savings mainly come from the revenues in the regulation market, and the cost for energy use and demand charge are almost the same in the *Base* and *OPT*. Because the sum of the RegD signal over a long time period (e.g. 1 hour) is almost 0, providing regulation service in the *OPT* leads to the similar energy use, thus similar energy cost compared with the *Base* where no regulation service is provided. By utilizing the demand cost defined in Eq. (25), the data center can provide regulation service without increasing monthly demand, thus no extra demand charge would be added to utility bills. The revenue from July is much higher than that in January because the price for dynamic regulation (RegD) resources is higher in July. As shown in Figure 4, during the studied two days, the average price from regulation market in July is about 21 \$/MW, while that in January is only about 5.8 \$/MW.

Figure 6 shows the hourly capacity bids in 1/21 and 7/21. The demand for each 30 minutes is denoted as the thin solid line. The demand limit used for demand cost as shown in Eq. (25) is denoted as the dashed line. The optimal capacity bid at each hour is denoted as the shaded area. At non-peak hours (e.g., 3:00 - 6:00), the optimal bid is mainly influenced by the price from energy market, price from regulation market and detailed shape of RegD signal. Because the demand is lower than the demand limit, the tradeoff between the energy cost and revenues from regulation market determines the optimal bid. The energy cost is highly influenced by the energy use, which is determined by the detailed shape of the RegD signal. If the sum of the RegD signal is larger than 0, then more energy would be consumed when providing frequency regulation service, thus the energy cost would increase. Although the energy cost increases in this case, the data center can get revenues from regulation market. If the sum of the RegD signal is no larger than 0, then at that hour, the data center can bid at their maximum capacity.

At peak hours (e.g., 12:00 - 16:00), the optimal bid is mostly influenced by the demand limit and the RegD signal. Figure 6 shows that at these hours, the bid is small so that the demand cannot exceed the required demand limit to avoid demand penalty. At 13:00, the bid is about 69 kW, but it is only about 5 kW at 14:00. The difference is caused by the detailed shapes of the RegD signals in these two hours. At 13:00, the sum of the RegD signal in first 30 minutes is slightly greater than 0, but in the second 30 minutes it is much smaller than 0. This means that regulation capacity bid in this hour can increase the demand in the first 30 minutes, but the demand in the second 30 minutes can be decreased compared with the same time in the baseline system. Therefore, at this hour, the data center can bid a large capacity as long as the demand in the first 30 minutes will not exceed the demand limit. The same situation happens at 14:00 but with a large sum of RegD signal at first 30 minutes. Also because the power at 14:00 is much closer to the demand limit, the data center can only bid a small capacity at this hour.

In summary, the proposed real-time optimization framework can help the data center without energy storage system harness the benefits from the energy market and the regulation market. However, the benefits are insignificant compared with the large baseline power in data centers. One of the reason is that data centers without energy storage system are difficult to limit their power demand during FR service, which contributes to a large portion of the utility bill. In the future, we will consider retrofit strategy (e.g., installing thermal storage energy system) in the data center to limit the power demand to maximize the benefit from the multi-markets

5 Conclusions

This paper developed a real-time multi-market optimization framework for the data center without storage systems to maximize their benefits from participating in both energy market and regulation market. Then, a case study was conducted to numerically investigate the optimal bids at each hour by considering the energy cost, demand costs and regulation revenues using a virtual data center located in PJM. Simulation results shows that using the proposed multi-market optimization framework can minimize the operational cost. Compared with the baseline system, providing frequency regulation service over the considered two days can save \$24.8 in January and \$123.6 in July.

6 Appendix

6.1 FR Performance Score

In the PJM market, new resources aiming to enter the regulation market need to pass an initial test by obtaining at least 0.75 for a defined performance score. The initial test signals of RegA and RegD are available at (PJM, 2019). The performance score is calculated as a composite score of accuracy, delay and precision, which are shown below (LLC, 2019).

$$c_{sig,res} = \frac{COV(reg,res)}{\sigma_{reg}\sigma_{res}} \tag{27}$$

$$S_{accuracy} = \max_{\delta=0-5\text{ min}} (c_{reg,res(\delta)}) \tag{28}$$

$$S_{delay} = \left| \frac{5\text{ min} - \delta^*}{5\text{ min}} \right| \tag{29}$$

$$S_{precision} = 1 - \frac{1}{n} \sum \left| \frac{res - reg}{\overline{reg}} \right| \tag{30}$$

$$S = \frac{S_{accuracy} + S_{delay} + S_{precision}}{3} \tag{31}$$

In the above equations, *reg* represents the regulation signal the DSRs receive from the electrical markets, and *res* represents the response signal the DSRs generate after control actions. c, COV and σ are the correlation coefficient, covariance, standard deviation of these two signals. In PJM, the response signal *res* is recalculated with a time shift δ ranging from 0 to 5 minutes in an increment of 10 seconds, which leads to 31 response signals $res(\delta)$. The accuracy score $S_{accuracy}$ is the maximum correlation coefficient c between *reg* and $res(\delta)$. The delay score S_{delay} is calculated based on the delay time δ^* when the maximum accuracy score is obtained using Eq. (29). The precision score $S_{precision}$ is defined as the relative difference between regulation signal and response signal, where n is the number of samples in the hour, and $\overline{reg}$ is the hourly average regulation signal. The final performance score S in that hour is calculated as the weighted average of the three individual scores.

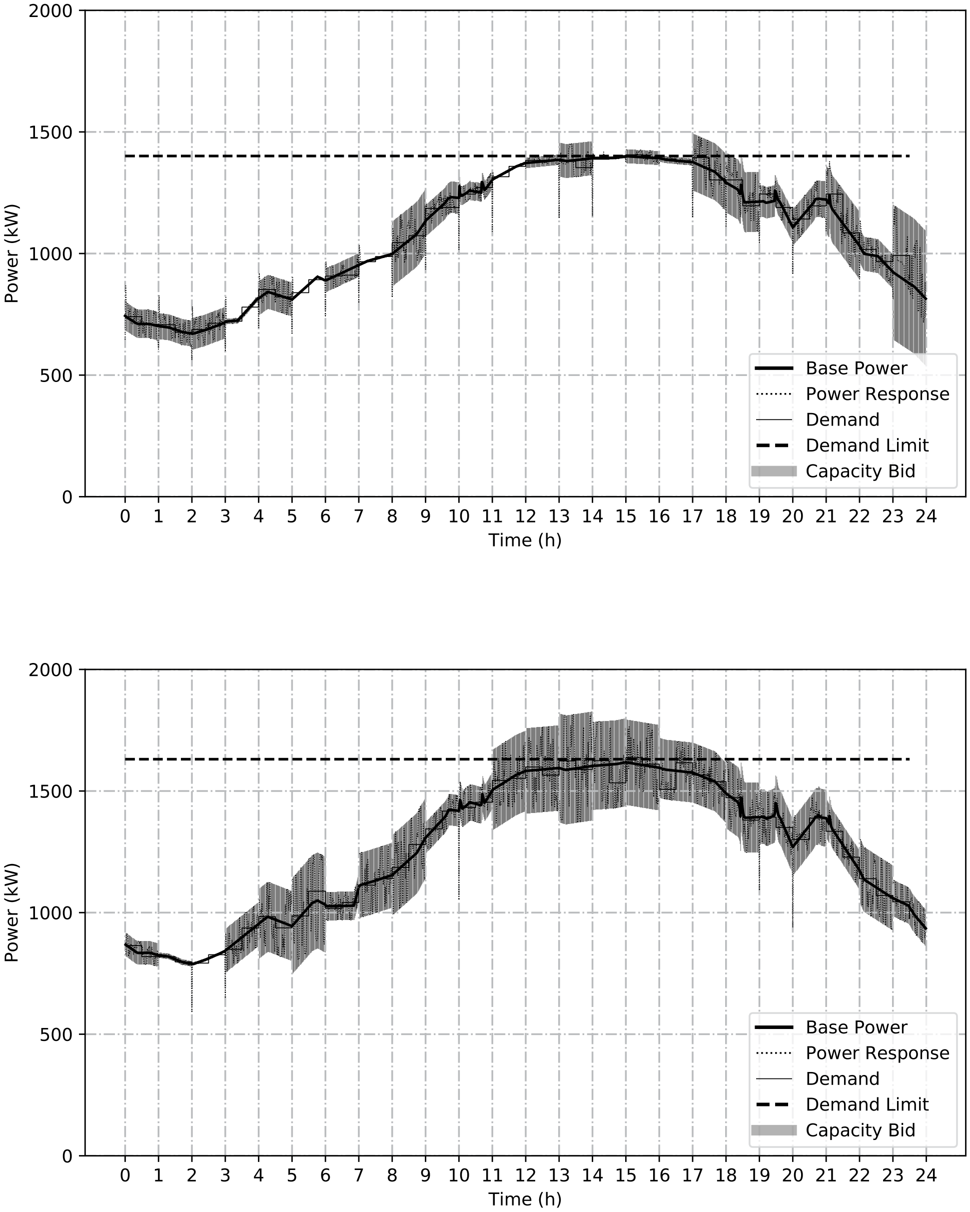

Figure 6. Optimal hourly regulation capacity bids in 1/21 (top) and 7/21 (bottom)

References

Natalie Bates, Girish Ghatikar, Ghaleb Abdulla, Gregory A Koenig, Sridutt Bhalachandra, Mehdi Sheikhalishahi, Tapasya Patki, Barry Rountree, and Stephen Poole. Electrical grid and supercomputing centers: An investigative analysis of emerging opportunities and challenges. *Informatik-Spektrum*, 38(2):111–127, 2015.

Gunter Bolch, Stefan Greiner, Hermann De Meer, and Kishor S Trivedi. *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications*. John Wiley & Sons, 2006.

Yangyang Fu, Michael Wetter, and Wangda Zuo. Modelica models for data center cooling systems. In *2018 Building Performance Analysis Conference and SimBuild, Chicago, Illinois, United States of America*, 2018.

Yangyang Fu, Xing Lu, and Wangda Zuo. Modelica models for the control evaluations of chilled water system with waterside economizer. In *Proceedings of the 13th International Modelica Conference, Regensburg, Germany, March 4ñ6, 2019*, number 157, page 8. Link^ping University Electronic Press, Link^pings universitet, 2019a.

Yangyang Fu, Wangda Zuo, Michael Wetter, James W. VanGilder, and Peilin Yang. Equation-based object-oriented modeling and simulation of data center cooling systems. *Energy and Buildings*, 198:503 – 519, 2019b. ISSN 0378-7788. doi:https://doi.org/10.1016/j.enbuild.2019.06.037. URL `http://www.sciencedirect.com/science/article/pii/S0378778819307078`.

Yangyang Fu, Wangda Zuo, Michael Wetter, Jim W. VanGilder, Xu Han, and David Plamondon. Equation-based object-oriented modeling and simulation for data center cooling: A case study. *Energy and Buildings*, 186:108 – 125, 2019c. ISSN 0378-7788. doi:https://doi.org/10.1016/j.enbuild.2019.01.018. URL `http://www.sciencedirect.com/science/article/pii/S0378778818330573`.

Sen Li, Marco Brocanelli, Wei Zhang, and Xiaorui Wang. Data center power control for frequency regulation. In *2013 IEEE Power & Energy Society General Meeting*, pages 1–5. IEEE, 2013.

PJM Interconnection LLC. Pjm manual 11: Energy & ancillary services market operations, 2019.

PJM. Ancillary services, 2019. URL `https://www.pjm.com/markets-and-operations/ancillary-services.aspx`.

Telecommunication Industry Association. Standards, Technology Dept, and American National Standards Institute. *Telecommunications Infrastructure Standard for Data Centers*. Telecommunication Industry Association, 2005.

Wei Wang, Amirali Abdolrashidi, Nanpeng Yu, and Daniel Wong. Frequency regulation service provision in data center with computational flexibility. *Applied Energy*, 251:113304, 2019.

Michael Wetter et al. Genopt-a generic optimization program. In *Seventh International IBPSA Conference, Rio de Janeiro*, pages 601–608, 2001.

Adam Wierman, Zhenhua Liu, Iris Liu, and Hamed Mohsenian-Rad. Opportunities and challenges for data center demand response. In *International Green Computing Conference*, pages 1–10. IEEE, 2014.