Efficient Renewable Generation through Load-Side Dynamic Reserve Provision

M. Caramanis, A. K. Coskun, I. Paschalidis Boston University; S. Reda, Brown University

Position Paper, 2013 NSF National Workshop on Energy Cyber-Physical-Systems

Abstract -- We advocate mission critical research for enabling electricity load-side-provision of reserves to support significant integration of volatile renewable generation. We claim that complete multi-scale electricity markets provide sufficient statistics of the power system, allowing distributed decision support cyber layers to interact with distributed load physical layers. Data centers, with their significant and growing share in electricity consumption and remarkable degrees of freedom, are used to elaborate the potential of load-side reserve provision.

I. CONTEXT

Power markets, introduced in the US in 1997, have been widely adopted, serving today the majority of high-voltage-connected generators and large consumers. The structure of these markets has already evolved to handle:

- From planning to operation decisions derived from cascaded markets clearing at time scales ranging from years (generation capacity and financial transmission rights markets), to months (commodity markets), to hours (day ahead -DA- and hour ahead -HA- markets), to five minutes (ex post marginal cost markets for the reconciliation of differences), to four seconds (regulation service -RS- signal update), and, finally, real time markets (frequency control).
- Co-optimization/co-clearing of energy and capacity reserves (primary for frequency control, secondary for regulation service, tertiary, etc.), whose system-level requirements reflect contingency planning for uncertainty in energy balance, transmission and supply availability.

Social-welfare contributions of competitive power markets are arguably due to the fact that they enable distributed, yet collaborative, decisions that (i) take advantage of local uncertainty and state dynamics information, and (ii) respond to price or other system-wide aggregators, such as frequency or the regulation signal, that act as sufficient statistics of the otherwise intractably complex system information. Whereas centrally controlled generation and demand control bear similarities to market-based distributed decision making, it suffers in two significant ways: first it lacks local dynamics and capacity information, and second, it cannot produce social-welfare-based economic valuation of consumer decision and risk preferences that are capable of guiding optimally operational as well as longer-term decisions. In conclusion, if we can guarantee complete markets—a task worth pursuing in itself, we can achieve the desired decoupling of distributed decision making from intractable system complexity. This decoupling enables both operational and longer term—investment, locational and the like- efficiencies. The rest of this position paper presents indicative mission critical research issues in Cyber-Physical-Electric-Energy-Systems related to complete electricity markets and harvesting the associated efficiencies of decoupled system complexity and distributed decision making.

II. MISSION CRITICAL RESEARCH ISSUES

II.1 Complete Markets

II.1.1 Multiple Period Markets: Flexible Storage-Like Loads and Renewable Generation Bids. Current day ahead (DA) multi-period markets co-optimizing energy and reserve transactions are deficient in two major ways: (*i*) they allow only myopic, price-quantity bids which are inadequate in representing a participant's multi-period utility, and (*ii*) do not charge renewable generation bids for the additional reserve requirements that their volatility imposes on the power system. These deficiencies provide incentives for storage-like loads such as HVAC, Electric Vehicle (EV) battery charging and storage resources to game the current bidding system leading to a Nash equilibrium that is only asymptotically socially optimal [1]. Similarly, renewable generation has the incentive to bid aggressively since it is not responsible for the reserve costs it imposes on the system [2,4]. Appropriate extension of bidding rules may remove gaming incentives under asymptotic market competitiveness conditions [1,2]. Markets spanning multiple periods with heterogeneous loads may also allow flexible reschedulable loads to extract a socially unfair advantage from inelastic loads [3]. *Mission critical research in completing multiple period markets* should focus on (*i*) extending bidding rules to allow the expression of flexible load inter-temporal utility and internalization of the impact of renewable generation on system reserve requirements, (*ii*) understanding and preventing non-competitive market situations, and (*iii*) addressing algorithm design and computational issues related to market clearing as they become more complex by the extended bidding rules and the associated inter-temporal constraint relationships.

II.1.2 Establishment of Distribution Markets. Distribution network marginal-cost-based prices of real and reactive power promotes full integration of high and low-voltage costs and enables retail market participants to provide much needed reserves, which, in essence, commoditize the quality of power supply and approach, asymptotically, the ideal of real-time-price demand response [5,12,15,17,19,29]. *Mission critical research in completing distribution markets* should focus on (*i*) linearization/ convexification of AC load flow equations [11,16] and (*ii*) developing parallel, distributed asynchronous optimization, and possibly cloud-based market clearing approaches enabling distribution markets with millions of participants [13,14].

II.2 Identification of Power System Requirements and Control of Transmission Systems Congestion

Power system requirements such as primary, secondary and tertiary reserves, as well as transmission line flow capacity constraints and other contingencies are critical quantities that allow markets to secure a stable and high-quality electricity supply. Determination of these quantities is a synthetic task involving both the physical power system layer capabilities (described by hybrid multi-time scale stochastic dynamics) as well as the cyber layer of information dissemination and the design of new proliferating types of capacity reserves. Although important, system requirements identification is out of the scope of this position paper and will not be discussed further. We will instead discuss mission-critical research issues associated with resilient transmission infrastructure operation, maintenance and planning.

II.2.1 *Transmission Topology Control* is an effective approach to decrease line flow congestion addressed today by optimizing generation commitment and economic dispatch. Recent research has shown [23, 30, and many other] that dynamic transmission line switching and FACTS device control can tractably decrease congestion costs to the tune of $200M per year, while making sure that line switching is robust to steady-state and transient stability implications of line switching. A natural extension of recent research results is to (*i*) combine generation unit commitment with *transmission line commitment*, (*ii*) redefine generation and transmission maintenance scheduling to reflect the optimal operation of transmission topology, and finally, (*iii*) redefine transmission and generation planning on the presupposition that optimal transmission topology control will be exercised during the maintenance and operation time scales. Given the expected proliferation of renewable generation, particularly considering the wind farms located far from load centers, increasing the resilience of existing transmission lines to new loads is of paramount importance.

II.3 Demand Response: System to Market Participant Interface Elaborated in the Case of Data Centers

Returning from system wide considerations, we focus next on optimal provision of *reserves* by flexible loads connected to the sub-transmission or distribution network. We assume full market participation privileges are available to loads with storage characteristics including, data centers and broadly construed computing services, HVAC plants with variable-speed-motor driven heat pumps, electric vehicle battery charging, duty cycle appliances and other existing or about to be introduced flexible loads. Demand response is associated with a basket of energy and reserve transactions. Specifically, a demand responsive market participant may purchase energy and sell reserves, thereby (*i*) reducing its effective cost of energy, while (ii) simultaneously reducing the quality of energy it receives by virtue of the fact that it has undertaken the obligation to provide reserves. The key difference relative to centrally administered demand response is that the voluntary decision to bid for reserves and to control its provision by responding to Independent System Operator (ISO) requests is less costly than it would have been under centrally controlled demand response. The enabler of efficient response to unpredictable ISO requests is the knowledge of their statistical properties. On the basis of these properties, a provider of reserves can decide on how much reserve to offer by estimating the expected optimal response cost.

Of the many types of demand responding loads that have been investigated [6, 7, 10, 19, 21, 27], we elaborate by focusing on data centers and their ability to offer regulation service (RS) reserves (also known as secondary reserves). The ability of data centers to offer RS is indeed significant due to their degrees of freedom in modulating their power consumption and the diversity of jobs that they process ranging from high priority transactional jobs to less sensitive jobs that require a reasonable processing rate on average rather than an immediate response. The selection of data centers is quite opportune given their increasing share in power consumption, which is 3% of total US electricity and growing. Interestingly, if one includes computing and communication infrastructure in buildings, or even dense urban areas (e.g., local computing facilities, infrastructure to support a cellular network), then, the percentage of total US energy consumption becomes larger. The focus on RS reserves is also interesting because of their higher value (i.e., their clearing price is comparable to the energy clearing price) relative to slower dynamics spinning or tertiary reserves and their more challenging modeling requirements. Finally, RS reserve requirements are increasing with renewable penetration [9].

II.3.1 Data Center Demand Response through RS Reserve Provision. Consider a data center that purchases in the hour-ahead market E MWh of energy at the clearing price Π^E, and sells secondary or regulation Service (RS) reserves $R<E$ at its clearing price Π^R. The net cost of this transaction to the Data Center is $E\Pi^E$-$R\Pi^R$. During the hour, the data center will have to observe the RS signal $y(t)$ and modulate its power consumption to track the implied obligation of $p(t)=E+y(t)R$. The RS signal $y(t)$ is an unanticipated decimal number in [-1,+1] broadcasted to all RS reserve providers at 4 second intervals, namely at t= 0,4,8,…,3596, 3600sec. The values of $y(t)$ are the output of an integral proportional filter of system frequency and balancing area control error, and as such they are unpredictable and independent of individual market participant behavior. The statistical behavior of $y(t)$ is well defined. The average value of the 3600/4=9000 $y(t)$ broadcasts over the hour is zero; i.e., the RS signal is energy neutral. The cost to the data center resulting from the operational level obligation to consume at the rate $p(t)=E+y(t)R$, consists of efficiency losses plus the value of reduced Quality of Service (QoS) offered to its clients during low $y(t)$ signal values.

An interesting, yet challenging, Cyber Physical Energy System emerges for a well-defined interface, $y(t)$, between a data center and the power system. The objective is to determine in the hour-ahead market the optimal level of average energy consumption E, and RS reserves offer, R, together with the associated optimal policy for responding to $y(t)$ during the hour that follows, subject to stochastic job arrivals, contractual QoS constraints and physical layer constraints and allowable actuations. We elaborate next by describing the cyber and physical layers and their interaction.

Cyber Layer. The cyber layer's objective is to optimize the following hybrid discrete event system. Given discrete probabilistic arrivals of processing requests (jobs) and a well-defined stochastic process describing the behavior of $y(t)$, determine a dynamic optimal control policy that maps the system state $x(t)$ to action $u(t)$ where:

- $x(t)$ contains (*i*) the state of servers in the data center –active, idle, asleep, off, in transition–, (*ii*) the jobs waiting in buffer queues or being processed, (*iii*) the current value of y(t), and (*iv*) the QoS achieved so far.
- $u(t)$ is a member of the allowable control set containing (*i*) initiation of server state transitions, (*ii*) assignment of jobs to servers and virtual machines, (*iii*) rerouting jobs to other data centers, and (*iv*) taking resource control and/or voltage and frequency scaling (DVFS) (or other power and performance management) actions at individual servers.

State dynamics depend on $u(t)$ and evolve with multiple time scale hybrid dynamics responding to discrete control actions (e.g.,a server state transitions) discrete events (e.g., a job arrival) and continuous retired instruction rates in response to DVFS settings.

Recent work [8, 28] on the cyber layer indicates that substantial cost reduction opportunities (e.g., 30% in preliminary results) exist when we regulate the computational power in accordance with ISO RS requests. After determining E and R levels based on workload estimates, physical limits of the servers, and constraints on performance and tracking error, it is possible to design a dynamic policy that will maintain the desired QoS level while tracking the ISO signal with a small error. This ability translates into cost savings (as explained in II.3.1). Components of the optimization problem (workload estimation, determining E and R, dynamic policy, etc.) individually and their interaction introduce interesting yet complex challenges. Understanding and exploiting synergies from the exchange of information between the cyber and physical layer is essential in resolving these challenges.

Physical Layer. Servers in data centers offer many degrees of freedom (DoF) that enable trading-off between performance and power consumption, which can enable effective regulation service while maintaining QoS for the users. Past work has investigated methods to determine the optimal server settings for the following DoF: sleep modes, dynamic voltage and frequency settings, number of active cores, and workload consolidation decisions [23, 24, 26]. Our recent work demonstrates that statistical profiling techniques can enable power/performance customization based on the characteristics of workloads and their interactions with each other [23]. Learning-based profiling methods are able to recognize emerging workload characteristics such as multi-threaded applications and virtual machines [24, 26]. Such learning methods enable adjusting the physical control knobs to maximize the application performance automatically while closely tracking a given power cap.

Furthermore, data center power optimization methods enable allocation of a large power budget among a cluster of servers such that the system's total performance is maximized while meeting the power budget. Given the intricate interactions between the computing servers and cooling equipment (e.g., CRAC units) in a data center, it is essential to provide self-consistent power budgeting, where the total power budget allocated between the computing servers and CRAC units ensure that the CRAC units are able to extract the heat of the servers while maintaining the server temperatures at reliable levels [25]. Self-consistent allocation enables better regulation by optimally allocating the power budget across the cyber components (i.e., servers) and the physical components (i.e., CRAC units) in a data center, while optimizing the total center's performance.

Cyber Physical Layer Interaction. Recent work on cyber layer decision support and physical layers optimization [8, 23, 24, 25, 26, 28] suggests fruitful directions for building efficient and robust interfaces between the two layers with the objective of (*i*) translating cyber layer action recommendations to implementable actuator settings, and (*ii*) feeding back to the cyber layer information obtained at the physical layer about DoF saturation limits, cooling requirement dynamics and Cooling Plant storage capabilities. This information can be used by the cyber layer for machine learning enabling it to dynamically revise/adapt its modeling assumptions.

III. CONCLUSION

This position paper makes the case that mission critical CPS research in (*i*) Electricity Market Design, and (*ii*) broadly construed demand response, can leverage the smart grid to marshal the cost effective reserves needed to support massive renewable generation and an energy sustainable computing sector.

REFERENCES

[1] M. C. Caramanis, E. Goldis, P. A. Ruiz, and A. Rudkevich, "Power Market Reform In the Presence of Flexible Schedulable Distributed Loads. New Bid Rules, Equilibrium and Tractability Issues" *Proceedings, 50th Annual Allerton Conf.* Oct. 1-5, 2012.

[2] M. C. Caramanis, J. M. Foster, "Uniform and Complex Bids for Demand Response and Wind Generation Scheduling in Multi-Period Linked Transmission and Distribution Markets", *Proc., 50th CDC*, 2011.

[3] Q. Huang, M. Roozbehani, and M. Dahleh, "Efficiency and Risk Impacts of Market Architecture in Power Systems under Realtime Pricing" *Proc., 51st CDC*, Dec 2012

[4] Varaiya, P.P., F.F. Wu, J.W. Bialek, "Smart Operation of Smart Grid Risk Limiting Dispatch" Pr. of the IEEE 99(1) pp. 40-57, 2011

[5] E. Ntakou M. C. Caramanis, "Price Discovery in Dynamic Power Markets with Low-Voltage Distribution-Network Participants", *Proc. IEEE PES TD conf.* April 2014.

[6] Enes Bilgin, M. C. Caramanis, "Decision Support for Offering Load-Side Regulation Service Reserves in Competitive Power Markets", *Proc., 52nd CDC*, December 2013

[7] E. Bilgin, M. C. Caramanis, Ioannis Ch. Paschalidis "Smart Building Real Time Pricing For Offering Load-Side *Regulation Service Reserves" Proc., 52nd CDC,* December 2013

[8] H. Chen, A. K. Coskun, and M. C. Caramanis "Real-Time Power Control of Data Centers for Providing Regulation Service" *Proc. 52nd CDC*, December 2013.

[9] Y.V. Makarov, C. Loutan, J. Ma, and P. de Mello, "Operational Impacts of Wind Generation on California Power Systems," *IEEE Trans. on PS*, vol. 24, no. 2, pp. 1039 - 1050, May 2009

[10] J. M. Foster, M. C. Caramanis "Optimal Power Market Participation of Plug-In Electric Vehicles Pooled by Distribution Feeder" *IEEE Trans. on PS*, August 2013

[11] J. Lavaei, S. Low, "Zero Duality Gap In Optimal Power Flow Problem", *IEEE Trans on PS*, V.27, No. 1, Feb. 2012.

[12] H. Chiang, M. E. Baran, "On the existence and uniqueness of load flow solution for radial distribution networks", *IEEE Trans on CS*, V37-No.3, March 1990.

[13] M. Kraning, E. Chu, J. Lavaei, and S. Boyd, "Dynamic Network Energy Management via Proximal Message Passing", *Foundation and Trends In Optimization*, to appear, 2013.

[14] A. D. Dominguez-Garcia, S. T. Cady, C. N. Hadjicostis, "Decentr. Opt. dispatch of distr. energy resources", *Trans. of 51st CDC*, 2012.

[15] J. Warrington, P. Goulart, S. Mariethoz, M. Morari, "A market mechanism for solving multi-period optimal power flow exactly on AC networks with mixed participants", *ACC 2012*, June 27- 29, 2012.

[16] J. Lavaei, S. Sojoudi, "Competitive Equilibria in Electricity Markets with Nonlinearities" *ACC 2012*, June 27-J29, 2012.

[17] Sathyanarayana, Heydt, "Sensitivity-based pricing and optimal storage utilization in distribution systems," IEEE Trans on Power Delivery, Vol. 28, No. 2, April 2013.

[18] R. Li, Q. Wu, S. S. Oren," Distribution Locational Marginal Pricing for Optimal Electric Vehicle Charging Management", *IEEE Trans on PS*, to appear.

[19] R.A.Verzijlbergh, Z.Lukszo, J.G.Slootweg, M.D.Ilic, "The impact of controlled electric vehicle charging on residential low voltage networks", 2011 International Conf on Networking, Sensing and Control.

[20] J. Wang, "A planning scheme for penetrating embedded generation in power distribution grids", PhD dissertation, M.I.T., September 2013.

[21] Y.C. Paschalidis, B. Li, M. C. Caramanis "Demand-Side Management for Regulation Service Provisioning through Internal Pricing" *IEEE Trans. on PS*, Vol. 27, No. 3, pp. 1531-1539, August, 2012.

[22] P. A. Ruiz, J. M. Foster,A. Rudkevich, and M. C. Caramanis, "Tractable Transmission Topology Control using Sensitivity Analysis" IEEE Trans. on PS, V27(3) pp. 1550-1559, 2012.

[23] R. Cochran, C. Hankendi, A. Coskun and S. Reda, "Identifying the Optimal Energy-Efficient Operating Points of Parallel Workloads," *ACM/IEEE Intern. Conf. on Computer-Aided Design*, pp. 608-615, 2011.

[24] S. Reda, R. Cochran, and A. Coskun, "Adaptive Power Capping for Servers with Multi--threaded Workloads," *IEEE Micro Journal*, V32(5), pp. 64-75, 2012.

[25] X. Zhan and S. Reda, "Techniques for Energy-Efficient Power Budgeting in Data Centers," in Design Automation Conference, 2013.

[26] C. Hankendi S. Reda, A. Coskun, "Cap: Adaptive Power Capping for Virtualized Servers," *IEEE International Symposium on Low-Power Electronics and Design*, pp. 415-420, 2013.

[27] J. M. Foster, G. Trevino, M. C. Caramanis, and M. Kuss "Plug-In Electric Vehicle Coordination with Renewable Generation: Theory and Application" *IEEE Systems Journal*, To appear.

[28] H. Chen, C. Hankendi, M. C. Caramanis, and A. K. Coskun "Dynamic Server Power Capping for Enabling Data Center Participation in Power Markets" *Trans. ICCAD'13 technical program*, to appear, 2013.

[29] F. Schweppe, M. Caramanis, R. Tabors, and R. Bohn, *Spot Pricing of Electricity*. Norwell, MA: Kluwer, 1988.

[30]K. W. Hedman, M. C. Ferris, R. P. O'Neill, E. B. Fisher, and S. S. Oren, "Co-optimization of generation unit commitment and transmission switching with n-1 reliability," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 1052–1063, May 2010.