

Datacenters as Controllable Load Resources in the Electricity Market

Rui Wang, Nagarajan Kandasamy, and Chika Nwankpa
 Dept. of Electrical and Computer Engineering
 Drexel University, Philadelphia, PA 19130, USA
 Email: {rui.wang, kandasamy, nwankpa}@drexel.edu

David R. Kaeli
 Dept. of Electrical and Computer Engineering
 Northeastern University, Boston, MA 02115, USA
 Email: kaeli@ece.neu.edu

Abstract—Datacenters, being major consumers of power, can play an important role in the efficient operation of electrical grids. This paper develops an optimization framework to allow datacenters to operate as controllable load resources within the demand dispatch regime, a demand response (DR) program in which incentives are designed to induce lower electricity use not just during times of high prices but also when the reliability of the local grid is jeopardized or when the electricity supply and demand are unbalanced. Assuming the availability of geographically distributed and virtualized datacenters situated in multiple regional electrical markets, the basic idea is to migrate the workload in the form of virtual machines (VMs) between these centers to maximize the expected payoff. The proposed framework addresses issues specific to the demand dispatch of datacenters such as timeliness of VM migrations and the impact of geographic distance on migration times. It also explicitly incorporates risks that may cause the load curtailment operation to be ultimately unsuccessful and result in monetary losses to datacenter operators; specifically, variability in network bandwidth that can cause uncertainty in VM migration times as well as the uncertain payoff when participating in DR markets. A set of case studies involving datacenters participating in an economic DR program is used to validate the framework.

Index Terms—Datacenters, demand response market, demand dispatch, risk-aware optimization.

I. INTRODUCTION

When electricity demand is close to the peak level supported by the power grid, this demand is usually satisfied using less efficient and higher cost generators; prices in wholesale markets can increase from about fifty dollars per MWh off peak to hundreds of dollars at peak hours. Under such circumstances, even a small reduction in demand can result in an appreciable reduction in system marginal costs of production. In recent years, *demand response (DR) programs* designed to encourage consumers to modify their electric demand level and usage pattern have become increasingly critical to the reliable operation and transmission efficiency of an (aging) electrical grid. Moreover, in deregulated electricity markets where the marginal generating unit determines the market price for all load, a drop in wholesale peak prices means that prices for all customers are also held in check—an important social benefit [1]. A specific type of DR program called *economic DR* is designed to incentivize customers to reduce power consumption in response to price signals from the utility when the market price, so called *locational marginal price (LMP)* of energy is high.

Datacenters are major consumers of electricity: in 2007, there were about 30 million servers worldwide, consuming 100 TWh of energy—almost 0.5% of the world's energy production—at a cost of \$9 billion; consumption increased even more so to 200 TWh in 2010 and has been projected to grow at about 9% annually [2], [3]. Therefore, datacenters can play an important role in ensuring the reliable and efficient operation of electrical grids worldwide while generating additional revenue from the electricity market.

This paper develops a risk-aware optimization framework that allows a datacenter to be treated as a DR resource: specifically, a *controllable load resource (CLR)* that can effectively participate in wholesale energy markets by reducing electricity consumption in response to control signals from the regional transmission operator (RTO). The DR resources are paid the prevailing LMP or some fraction thereof for such reductions. Assuming the availability of geographically distributed datacenters situated in multiple regional electrical markets, the basic idea is to migrate the incoming workload appropriately between these centers to maximize the expected payoff while accounting for the inherent monetary risks involved in doing so to the datacenter operators. This idea follows from the following observations:

- Global companies operate multiple datacenters hosting thousands of servers; for example, Citibank operates an estimated fourteen datacenters worldwide whereas the HSBC banking group operates eight datacenters comprising about 15,000 servers; HP operates fifty datacenters, mostly in the US.¹
- The average utilization of servers in datacenters is low: a recent study by Gartner, Inc, in 2010 found that the average server utilization ranged between 15-20%.
- Current and predicted network delays per request will not change significantly when computations move from one electrical market region to another.
- Virtualization is increasingly used in datacenters to consolidate workload and improve server utilization. This

¹The estimates were obtained from *Data Center Knowledge*, a news and analysis site about the datacenter industry that tracks the growth of cloud computing and datacenter technologies (www.datacenterknowledge.com). Other obvious examples of large datacenters are those operated by companies such as Amazon, Google, Yahoo, and eBay; for example, it is estimated that Amazon operates about seven datacenters comprising 450,000 total servers.

technology enables a single server to be shared among multiple performance-isolated platforms called virtual machines (VMs), where each VM can support one or more services. The disk space, memory, and CPU share allotted to VMs can be dynamically tuned; and more importantly for our purposes, live VMs can be migrated between hosts at run time while preserving the state of active user sessions. Current virtualization solutions allow VMs to be migrated between datacenters situated as far away as 500 km (see Section II).

The foregoing discussion implies that organizations operating multiple datacenters have the spare capacity and the flexibility needed to dynamically distribute workload between these centers. Assuming Citibank as an example, the company can register its various datacenters as CLR's in their respective electricity markets; an individual CLR can then respond to dispatch signals from the RTO to reduce (electricity) load by migrating some of the incoming workload in the form of VMs to other datacenters. Here, VMs can be migrated in two possible ways: along with the data or just the computation, leaving the data at the source site and accessing it remotely. Also, note that, modulo security, workload migration need not be confined to datacenters belonging to a single organization; Citibank could purchase computing capacity on an external cloud service such as Amazon's EC2 and move portions of the workload there, thereby relieving load on its datacenters.

Prior research has developed load-balancing techniques for geographically distributed datacenters that take into account the prevailing LMP prices (discussed in more detail in Section V). The basic idea is one of self-scheduling: redistribute the workload between datacenters, thereby reducing the load of datacenters situated in areas with high LMPs. This is a purely voluntary action initiated by the customer to reduce the consumption of electric energy from the expected consumption in response to an increase in the price of electric energy.² The fundamental difference between the framework described here and existing work is that our techniques allow datacenters to operate as CLR's within the *demand dispatch regime*, a more sophisticated form of DR in which incentives are designed to induce lower electricity use not just during times of high LMPs but also when the reliability of the local grid is jeopardized or when the electricity supply and demand are not in balance [5]. DR can now occur at any time during the day and not just during high LMP periods. In addition to the traditional method of dispatching generators to raise/lower electricity production, utilities can also dispatch CLR's to increase/decrease their load by a specific amount; these CLR responses produce the same results as changing the output of generating plants. So, demand dispatch is the complement of generation dispatch in that a CLR behaves as a *virtual power plant*; for example, if a CLR

is dispatched to drop 100 kW in load, this is equivalent to an increase in generation capacity by 100 kW when viewed at the tie-line level of the grid. Demand dispatch has the potential to efficiently mitigate local overload conditions as they happen while reducing the electrical losses associated with transporting electric energy from remote generation sources.

This paper develops an optimization framework to allow datacenters to participate as CLR's within the demand dispatch regime with the overall goal of maximizing the expected monetary payoff. Our solution addresses the following key issues:

- Utilities currently issue DR notifications with a lead time of two to three hours along with requirements on the amount of load to be curtailed, usually ≥ 100 kW. So, datacenters wishing to participate must coordinate and complete their VM migrations, and curtail the load within this deadline. The migration time incurred per VM is a function of the available network bandwidth and the network latency between datacenters, limiting the number of VMs that can be migrated between two centers within the deadline. This constraint is explicitly considered in the optimization problem.
- Once the DR resource is dispatched by the utility, failure to achieve the said load reduction when participating in certain DR programs incurs stiff financial penalties. So, when making a decision to commit to such DR programs, the optimizer must consider the various risks that may cause the curtailment operation to be ultimately unsuccessful. One such risk factor is when VMs are in the process of being migrated during the two hour window but fluctuations in the available network bandwidth cause migration times to increase more than anticipated, and the number of migrated VMs is insufficient to achieve the promised load reduction. This uncertainty/risk is incorporated explicitly within the optimization framework as well.
- Finally, there is the issue of the financial reward: once the resource is dispatched by the RTO, the actual reward is only known after the time of the settlement. So, the optimizer must continually learn the outcome of its DR actions and use it to guide future decisions.

We present a case study involving a set of datacenters participating in an economic DR program to validate the proposed framework. Simulation results show that the optimizer affects a reasonable tradeoff between the number of VMs migrated over a link and the bandwidth fluctuation on that link. Also, the migration scheme is risk-averse to the variance in electricity prices and aims to achieve a tradeoff between the expected profit and the variance associated with the profit.

The paper is organized as follows. Section II familiarizes the reader with the demand dispatch paradigm in the US electricity market. Section III develops the optimization framework and Section IV presents a case study using a set of distributed data centers. Section V discusses related work and we conclude the paper in Section VI.

²A recent order implemented in April 2012 by the Federal Energy Regulatory Commission (that has jurisdiction over inter-state and wholesale electric rates) has eliminated the option for self scheduling and DR resources must now be dispatched by the responsible utility; that is, DR refers only to energy modifying activities performed in response to explicit economic/reliability signals provided by the utility [4].

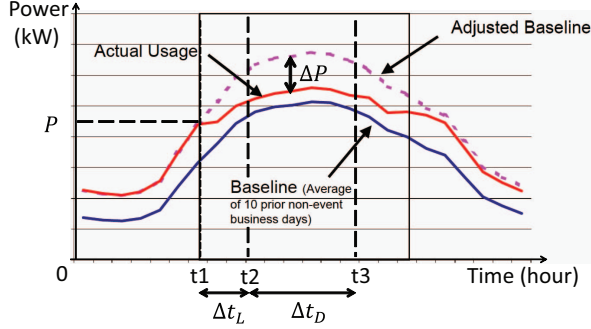


Fig. 2. The operating principle underlying the RTDR and EDR programs, summarized from [8]. Here, Δt_L and Δt_D denote the lead time and the downtime, respectively, and $\Delta \bar{P}$ denotes the load curtailment requested by the RTO. The actual measured load curtailment is denoted by ΔP .

differences on that particular day). The minimum $\Delta \bar{P}$ that the datacenter must achieve varies by RTO and typically lies in the 50-100 kW range. The deviation between the requested reduction and the actually measured curtailment should be less than 20%, and the notification lead time and downtime are usually a couple of hours [4].

In the RTDR program, the reward R during a given hour t is based on the measured curtailment ΔP and the real time LMP_t as (1) shows

$$R_t = \begin{cases} \Delta P \cdot LMP_t & LMP_t \geq NBP \\ \Delta P \cdot (LMP_t - GT) & GT \leq LMP_t < NBP \\ 0 & LMP_t \leq GT \end{cases} \quad (1)$$

where GT is the generation and transmission components of the participant's electric bill [4]. In the EDR program, R_t is

$$R_t = \begin{cases} \Delta P \cdot \max(\$500/MWh, LMP_t) & \text{voluntary} \\ \Delta P \cdot \max(\text{StrikePrice}, LMP_t) & \text{mandatory} \end{cases} \quad (2)$$

where the strike price (up to \$500/MWh) is the price at which the customer is willing to participate in the market [1]. As we can see from (1) and (2), the reward depends on the prevailing LMP values from t_2 to t_3 (when the demand is actually curtailed). Since these LMP values are unknown at t_1 , participants must make curtailment decisions based on estimated LMP values over the interval $[t_2, t_3]$.

We return to Fig. 2 to make one last point regarding the minimum granularity requirement on $\Delta \bar{P}$, the load curtailment requested by the RTO. It varies between RTOs: PJM, for example, sets the requirement at 100 kW whereas the New York ISO sets it at 50 kW. This requirement allows the RTO to monitor the participation of the CLR, that is, distinguish between load curtailment that is actually happening versus normal load fluctuation in the grid. The granularity requirements exist, in large part, due to limitations of current metering technology. As more advanced and sensitive metering infrastructure is deployed on the grid and the number of metering points increased, we expect the minimum curtailment granularity to decrease.

B. VM Migration Across Data Centers

The increasing use of dynamic power management schemes within a datacenter can mean that its real-time power consumption P is already (near) minimal, leaving little margin for further power curtailment (in terms of ΔP in Fig. 2) and precluding a single datacenter from responding to DR signals, especially given the minimum requirement on $\Delta \bar{P}$. However, if an organization operates a number of geographically distributed datacenters, then by migrating the workload appropriately (in the form of live VM migrations), from source application servers in one datacenter to destination application servers in other datacenters, we can shut down the source servers to trade ΔP for the payoff while still maintaining the quality of service (QoS). The VMs can migrate back to source servers after time t_3 . Here, we will assume that VMs within the application tier process session-based workload. So, to maintain the sessions of re-dispatched workloads, each VM's state must be copied over to the destination, while its storage (maintained within the database tier) can stay within the source datacenter. (For sessionless workloads, the corresponding VMs need not to be migrated to the destination; the necessary VMs can be freshly instantiated.)

In terms of VM migration over wide area networks, VMware's vMotion technology [9] combined with solutions from other vendors such as Cisco [10], F5 [11], and NetEx [12] allow VMs to be migrated between datacenters as long as the network satisfies minimum bandwidth (622 Mbps) and round-trip latency requirements (for example, not to exceed 10 milliseconds in vMotion 5.0, corresponding roughly to a one-way distance of 500 kilometers) [13]. The migration time for a VM varies depending on its memory usage [14] as well as the distance (network latency) and the available bandwidth between the datacenters; for example, published data shows that it takes about 80 seconds for a VM with 8 GB of memory to migrate between two datacenters 100 kilometers away [10]. During the migration process, the application hosted within the VM maintains all its sessions on the source server except for a transient drop in performance (again, measured to be around 10%) when the VM switches to execute on the destination server [10]. Finally, as all the VMs within a source server are migrated to a corresponding destination server, we must ensure that the CPU and memory capacity available at the destination matches the source, so that the current resource provisioning to the VMs can still be satisfied at the destination to maintain a steady QoS.

III. THE OPTIMIZATION FRAMEWORK

This section discusses the system modeling assumptions and describes how to pose the demand dispatch of datacenters as a risk-aware optimization problem.

A. System Modeling Assumptions

We assume the availability of geographically distributed datacenters situated in multiple regional electrical markets and inter-connected via high-speed communication links. Fig. 3 shows the various datacenter locations used in our case study.

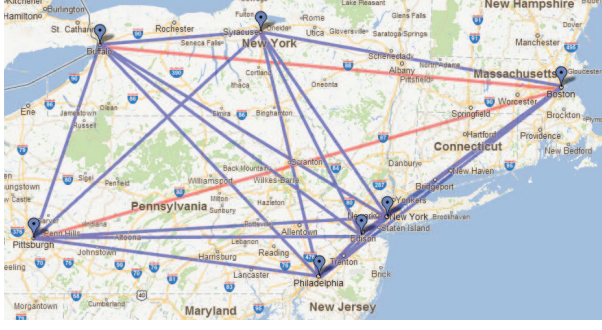


Fig. 3. The locations of distributed datacenters used in our case study. Blue lines indicate communication links less than 500 km between datacenters whereas red lines indicate longer links; live VM migration is not feasible over these links since we assume that VMs can only be migrated between datacenters within 500 km of each other.

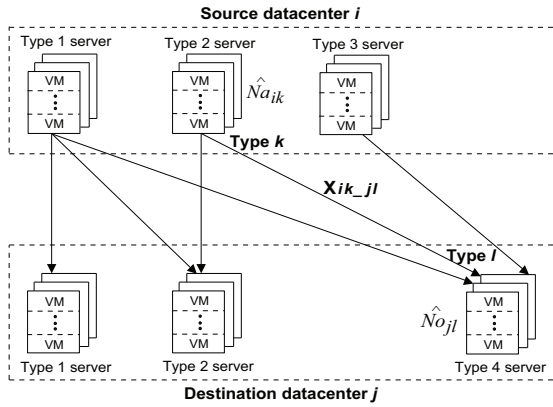


Fig. 4. The VM migration scheme from source datacenter i to destination datacenter j .

Application servers within a datacenter are categorized into multiple types based on the CPU and memory capacity, as shown in Fig. 4. A datacenter houses a mix of different server types and servers with more CPU/memory resources are denoted using larger type numbers. We use i to denote a source datacenter that migrates VMs out whereas j denotes a destination center accommodating those VMs, and $i \neq j$. If k and l denote server types within datacenters i and j , respectively, then we use x_{ik_jl} to indicate the number of type k servers in datacenter i migrated to type l servers within center j , where $l \geq k$ ensures that the destination server possesses more CPU/memory resources than the source server.

B. Formulation of the Basic Optimization Problem

We first formulate the basic optimization problem and then extend the formulation to incorporate risk/uncertainty in the following section.

If p_k represents the power consumed by a type k server, then

$$\Delta P_i = PUE_i \cdot \sum_k \sum_j \sum_l (p_k \cdot x_{ik_jl}) \quad (3)$$

represents the total amount of power curtailed at the source datacenter i by migrating out the VMs and shutting down

x servers. The energy efficiency of a particular datacenter is captured via its power usage effectiveness (PUE) metric which is the ratio of the total amount of power consumed by a datacenter (including the computing power as well as the corresponding cooling and infrastructure power) to the power consumed by just the computing infrastructure. Then, the expected monetary reward obtained by all source datacenters during time period Δt_D is

$$f_R(\mathbf{x}) = \sum_i \left(\sum_t \widehat{LMP}_{it} \right) \cdot \Delta P_i, \quad (4)$$

where $\widehat{LMP}_{it}$ is the estimated LMP value during hour t at datacenter i .

Migrating VMs as part of the load curtailment event may incur a round-trip transient cost which can be modeled as

$$f_C(\mathbf{x}) = 2 \cdot \sum_i \sum_k (c_k \sum_j \sum_l x_{ik_jl}), \quad (5)$$

where $2 \cdot c_k$ accounts for the fact that during the round-trip migration process, a fraction of the requests directed at a VM may violate the QoS requirement.⁵

Another term to be considered is the difference in electricity cost incurred due to the VMs that are migrated. The cost of operating a set of servers at the source datacenter may differ from the cost incurred for the same set of servers at the destination due to these reasons: (1) VMs from a source server may be accommodated by a destination server that consumes more power, (2) the PUE ratings of the source and destination datacenters may be different, and (3) the electricity prices may be different for the datacenters in their respective regional markets. Meanwhile, the total time spent by the x migrated servers at the destination datacenter comprises three parts: (1) half the lead time, on average, when in the process of moving into the destination, (2) the downtime, and (3) half the lead time when in the process of leaving the destination. The difference of electricity cost is then obtained as

$$f_E(\mathbf{x}) = \sum_i \sum_k \sum_j \sum_l \left[\left(\frac{\Delta t_L}{2} + \Delta t_D + \frac{\Delta t_L}{2} \right) \cdot (r_j \cdot PUE_j \cdot p_l - r_i \cdot PUE_i \cdot p_k) \cdot x_{ik_jl} \right]. \quad (6)$$

The above equation assumes that the datacenters purchase electricity at a flat retail price r in \$/MWh from their respective regional utilities. (If instead of the flat rate, the datacenters pay the real-time LMP, also in \$/MWh, for the power consumed, then r can simply be replaced by the estimated LMP values.)

The profit obtained by committing to a DR signal is the difference between the expected reward and the various associated costs:

$$f_P(\mathbf{x}) = f_R - f_C - f_E. \quad (7)$$

⁵Experiments conducted on a small-scale virtualized cluster indicate that when migrating a VM between two servers, a small fraction—between 5 to 10% of the requests directed at this VM experience significantly increased response times. The reader is referred to Kusic et al. for more details regarding the actual experimental setup and results [15]. The Cisco document on workload mobility reports similar results when migrating VMs between datacenters [10].

The final step is to enhance the above cost function to consider the impact on application QoS when only the VM's computational state is migrated over to the destination, leaving the storage within the source datacenter to be remotely accessed over the network. We achieve this by adding a distance factor f_D to (7) such that whenever possible, the optimizer prefers solutions that migrate VMs such that the computation remains geographically close to the corresponding data. The distance factor is mathematically stated as

$$f_D(\mathbf{x}) = \sum_i \sum_j (d_{ij} \cdot \sum_k \sum_l x_{ik_jl}), \quad (8)$$

where d_{ij} is the distance between source datacenter i and destination datacenter j . The final objective function can then be written as a combination of (7) and (8) as

$$\max_{x_{ik_jl}} f_P - \mu_D \cdot f_D \quad (9)$$

subject to

$$x_{ik_jl} \in \mathbb{Z}_{\geq 0}, \quad (10)$$

$$\sum_j \sum_l x_{ik_jl} \leq \widehat{N}a_{ik}, \quad (11)$$

$$\sum_i \sum_k x_{ik_jl} \leq \widehat{N}o_{jl}, \quad (12)$$

$$g_{Mij}(\mathbf{x}) \leq \Delta t_L, \quad (13)$$

$$0.8 \cdot \Delta \bar{P}_i \leq \Delta P_i \leq 1.2 \cdot \Delta \bar{P}_i. \quad (14)$$

Here μ_D is a weighting factor that affects the tradeoff between f_P and f_D . Constraint (11) ensures that the total number of type k servers migrated out of datacenter i does not exceed $\widehat{N}a_{ik}$, the estimated number of type k servers in datacenter i that should be operational to satisfy the incoming workload intensity. Similarly, (12) guarantees that the total number of servers migrated into type l servers in datacenter j does not exceed $\widehat{N}o_{jl}$, the estimated number of type l servers in datacenter j that are idle (meaning they can accommodate incoming VMs).⁶

The term $g_{Mij}(\mathbf{x})$ in (13) captures the total VM migration time spent on the link between datacenters i and j , and is calculated as

$$g_{Mij}(\mathbf{x}) = t_{ij} \cdot \sum_k (m_k \cdot \sum_l x_{ik_jl}), \quad (15)$$

where m_k is the memory capacity of a type k server, and t_{ij} is the time needed to migrate one GB of VM memory from datacenter i to j . Fig. 5, summarized from a Cisco report on virtualized workload mobility [13], shows that generally speaking, the migration time incurred per VM is a function of the geographical distance between two datacenters. Based on this observation, we model the migration time as

$$t_{ij} = t_0 + a \cdot \frac{d_{ij}}{w_{ij}}, \quad (16)$$

⁶Using the historical and current workload intensity, we can estimate both $\widehat{N}a$, the number of servers will stay active as well as $\widehat{N}o$, the number of servers that will stay idle during the entire curtailment event.

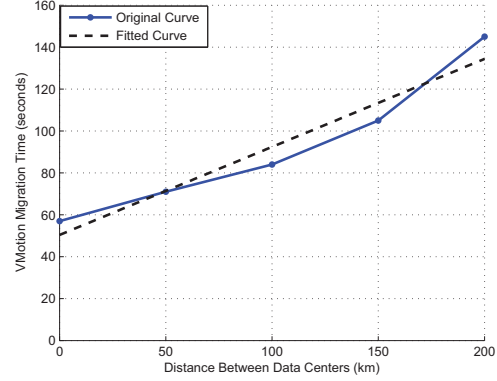


Fig. 5. The time incurred in migrating a VM as a function of distance between two datacenters. The VM used in this experiment hosted Microsoft SQL server and had a memory usage of 8 GB. The plot shows both the original curve as well as the best-fit linear line.

where t_0 is the base time, a is a constant factor, and d_{ij} and w_{ij} are the distance and bandwidth between the source and destination datacenters, respectively. Both t_0 and a can be easily estimated via a curve-fitting process. As per (13), migration of the chosen VMs must be completed within the available lead time Δt_L , which limits the number of VMs that can be migrated within the deadline per link. Finally, (14) guarantees that the actual power eventually curtailed by each source datacenter ΔP_i is within the 20% deviation of the required curtailment $\Delta \bar{P}_i$.

C. Modeling and Optimization of Risks

There are always risks to the datacenter operator when committing to a DR signal. For example, the available bandwidth between datacenters can dynamically fluctuate around the expected bandwidth at any given time and as a result, the total time taken to migrate VMs on a link maybe longer than expected, violating (13) and leading to the migration risk. In addition, the DR reward depends on the future LMPs during the downtime hours. Even if we can estimate their future values, the actual LMPs may be different. So the actual payoff obtained may deviate from the expected payoff, introducing a payoff risk.

In the modern theory of risk management, Markowitz's mean-variance (MV) portfolio approach analyzes a decision under uncertainty and evaluates it in terms of the tradeoff between the decision's payoff and risk [16] [17]. The payoff of a portfolio is the expected value of a certain return and the risk can be identified as the variance of the return. Given a return function $f(\mathbf{x}, \xi)$ that is dependent on the decision vector $\mathbf{x}$ and random element vector ξ (uncertainty), let $E[f(\mathbf{x}, \xi)]$ represent the measure of payoff and $\sigma^2(f(\mathbf{x}, \xi))$ represent the measure of risk.⁷ Then the MV model can be stated as the problem of minimizing the risk while requiring the payoff to

⁷Usually it is appropriate to think of ξ as being governed by a probability distribution that is known or can be estimated.

exceed a predefined threshold r_0 , as

$$\min_{\mathbf{x}} \{ \sigma^2(f(\mathbf{x}, \xi)) \mid E[f(\mathbf{x}, \xi)] \geq r_0 \}. \quad (17)$$

Alternatively, the MV model can also be formulated to maximize a weighted combination of the payoff and risk, as

$$\max_{\mathbf{x}} \{ E[f(\mathbf{x}, \xi)] - \mu \cdot \sigma^2(f(\mathbf{x}, \xi)) \mid \mu \geq 0 \}, \quad (18)$$

where μ is the weighting factor.

In terms of the bandwidth and LMP values, we can estimate them based on historical observations using well-known forecasting methods (for example, Holt's linear exponential smoothing [18]) and the resulting value V follows a normal distribution as per

$$V \sim N(E[V], \sigma^2(V)). \quad (19)$$

Migration risk: The total migration risk incurred due to uncertain bandwidth, denoted by vector $\mathbf{w}$, is written in the form of (17) as

$$\min_{x_{ik-jl}} \sum_i \sum_j \sigma^2(g_{Mij}(\mathbf{x}, \mathbf{w})) \quad (20)$$

subject to

$$E[g_{Mij}(\mathbf{x}, \mathbf{w})] \leq \Delta t_L, \quad (21)$$

where

$$\sigma^2(g_{Mij}) = (a \cdot d_{ij})^2 \cdot \sigma^2\left(\frac{1}{w_{ij}}\right) \cdot \left(\sum_k m_k \sum_l x_{ik-jl}\right)^2. \quad (22)$$

The value of $\sigma^2(1/w_{ij})$ increases with $\sigma^2(w_{ij})$ and can be calculated using the estimated $E[w_{ij}]$ and $\sigma^2(w_{ij})$ values.

Reward risk: The reward risk incurred due to the uncertain LMP values can be written in the form of (18) as

$$\max_{x_{ik-jl}} E[f_R(\mathbf{x}, \widehat{\mathbf{LMP}})] - \mu_R \cdot \sigma^2(f_R(\mathbf{x}, \widehat{\mathbf{LMP}})), \quad (23)$$

where $\widehat{\mathbf{LMP}}$ is the estimate for the LMP values. Also,

$$\sigma^2(f_R) = \sum_i [\sigma^2(\sum_t \widehat{LMP}_{it}) \cdot \Delta P_i^2], \quad (24)$$

and $\sigma^2(\sum_t \widehat{LMP}_{it})$ can be calculated via the estimated $E[\widehat{LMP}_{it}]$ and $\sigma^2(\widehat{LMP}_{it})$ values.

So, the profit function in (7) can now be modified as

$$E[f_P(\mathbf{x}, \widehat{\mathbf{LMP}})] = E[f_R] - f_C - f_E, \quad (25)$$

and the variance of profit is quantitatively the same as the variance of the reward:

$$\sigma^2(f_P(\mathbf{x}, \widehat{\mathbf{LMP}})) = \sigma^2(f_R). \quad (26)$$

As the final objective is to affect a tradeoff between the expected payoff and the above described risks, the improved optimization problem is now the balance of (8), (20), (25), and (26), as shown below:

$$\max_{x_{ik-jl}} E[f_P] - \mu_D \cdot f_D - \mu_M \cdot \sum_i \sum_j \sigma^2(g_{Mij}) - \mu_R \cdot \sigma^2(f_P) \quad (27)$$

subject to constraints (10), (11), (12), (14), and (21). Here, μ_D , μ_M , and μ_R are all weighting factors.

TABLE I
DISTANCES IN KM BETWEEN THE SEVEN DATACENTERS USED IN THE CASE STUDY.

Datcenter	1	2	3	4	5	6	7
1 Philadelphia, PA	0	81	128	433	348	444	412
2 New Brunswick, NJ	81	0	44	349	315	449	464
3 New York, NY	128	44	0	305	314	466	496
4 Boston, MA	433	349	305	0	417	634	764
5 Syracuse, NY	348	315	314	417	0	218	428
6 Buffalo, NY	444	449	466	634	218	0	283
7 Pittsburgh, PA	412	464	496	764	428	283	0

TABLE II
THE PUE RATINGS OF EACH DATACENTER AND THE RETAIL ELECTRICITY PRICES IN CENTS/KWH.

Datcenter	1	2	3	4	5	6	7
PUE	1.9	1.5	1.3	1.7	1.6	1.8	2.0
Retail Electricity Price	7.2	10.6	6.7	13.1	6.7	6.7	7.2

TABLE III
THE FOUR TYPES OF SERVERS HOUSED WITHIN THE DATACENTERS.

Server Type	1	2	3	4
Frequency $\times$ Cores (GHz)	8	10.8	16	22.4
Memory (GB)	4	4	8	8
Power Consumption (Watt)	160	220	320	440
Migration Cost (Cent)	0.8	1.1	1.6	2.2

D. Computation of Success Probability

Once the optimal VM migration solution $\mathbf{x}^*$ is obtained, we combine this solution with the estimation for **LMP** to obtain the expected profit $E[f_P^*]$ and the associated variance $\sigma^2(f_P^*)$. The datacenter operator can then decide whether or not to commit to the DR signal based on either of the following evaluations.

- The operator can calculate a threshold value, f_{p0} , for the expected profit that is guaranteed to be met or exceeded with some high probability $Prob_0$ (specified by the operator). In other words, $Prob(f_P^* \geq f_{p0}) = Prob_0$.
- The operator can evaluate the probability that the profit f_P^* exceeds a threshold f_{p0} ; that is, $Prob(f_P^* \geq f_{p0})$.

IV. CASE STUDIES

This section uses a set of case studies to validate the optimization framework. We first describe the key simulation parameters which reflect the situation in the real world. Then test scenarios are described to compare the results provided by the basic formulation with those provided by the risk-aware formulations. Scenarios in which multiple datacenters receive DR signals from their respective RTOs at the same time are also discussed.

A. Evaluation Parameters

Table I shows the seven datacenters assumed in the paper. As discussed in Section II, VM migration cannot happen between two locations more than 500 km away; when the distance is less than 500 km, the migration time is captured by (15). After curve fitting the experimental data on migration times reported in [10], t_0 in (16) is set to 6.3 and a is set to 0.0525. As shown in Table II, the datacenters also have

TABLE IV
NUMBER OF SERVERS IN EACH OF THE SEVEN DATACENTERS.

Data Center	Server Type				Total Number
	1	2	3	4	
1	600	800	600		2000
2	800	400			1200
3	700	1000	1000	500	3200
4		500		500	1000
5		400	200	400	1000
6	300		200		500
7	400		200	200	800

TABLE V
THE EXPECTED BANDWIDTH $E[w]$ (GBPS), STANDARD DEVIATION $\sigma(w)$ (%), AND $\sigma(1/w)$ (1/GBPS) FROM DATACENTER 1 TO OTHER DATACENTERS.

Datcenter	2	3	4	5	6	7
$E[w]$	2.49	2.49	1.24	2.49	1.24	1.24
$\sigma(w)$	10%	10%	10%	10%	10%	20%
$\sigma(1/w)$	0.041	0.041	0.084	0.041	0.084	0.199

TABLE VI
THE EXPECTED LMPs $E[LMP_{1t}]$ (\$/MWh) AND THE STANDARD DEVIATION $\sigma(LMP_{1t})$ (\$/MWh) FROM HOUR 15 TO HOUR 20 AT DATACENTER 1.

Hour	15	16	17	18	19	20
$E[LMP_{1t}]$	145	220	170	280	325	190
$\sigma(LMP_{1t})$	43.5	66.0	51.0	84.0	97.5	57.0

different PUE ratings. Also, each state in the United States charges a different retail electricity price to industrial customers; we use the data reported by the US Energy Information Administration [19].

Our simulations use the four different server types assumed in Table III, where each type possesses different CPU and memory capacity and consumes a different amount of power. We also assume that the transient migration cost incurred by a server is proportional to its processing capacity, though in the case of an actual system, this cost can be readily obtained via profiling the VMs. Each datacenter houses hundreds of servers with a mix of all four server types, as assumed in Table IV.

B. Evaluating the Basic Optimization Framework

In the first scenario, we evaluate the basic formulation developed in Section III-B where the uncertainties in bandwidth and LMP values are not considered (although they do exist). We focus on datacenter 1 and assume that it is connected to other datacenters via a SONET OC-24 (offering a bandwidth of 1244 Mbps) or OC-48 (offering a bandwidth of 2488 Mbps) link. Table V lists the expected bandwidth of the link as well as the variance of this bandwidth, which for the purposes of this case study was set to 10% on some links and 20% on others.⁸ Returning to Fig. 1, assume that datacenter 1 receives a RTDR signal at hour 12 from PJM requesting a power curtailment of 350 kW from hour 15 to the end of hour 20.⁹ Based on

⁸In a practical implementation, the bandwidth will dynamically fluctuate at run time and needs to be periodically measured and updated.

⁹PJM is the RTO responsible for the Philadelphia area whereas the Boston area, considered in the later scenario, is managed by ISO New England.

TABLE VII
THE ESTIMATED NUMBER OF ACTIVE AND IDLE SERVERS IN THE DATACENTERS.

Data Center	Estimated Active Servers				Estimated Idle Servers			
	1	2	3	4	1	2	3	4
1	570	800	560		30	0	40	
2	680	120			120	280		
3	700	690	1000	500	0	310	0	0
4		430		500		70		0
5		190	200	400		210	0	0
6	290		140		10		60	
7	400		140	200	0		60	0

the historical LMPs, the datacenter estimates the LMP values during the required downtime, as shown in Table VI.

Meanwhile, based on their own historical workload intensity, the seven datacenters independently and simultaneously estimate the number of servers that will stay active ($\hat{N}_a$) as well as idle ($\hat{N}_o$) during this event, an example of which is shown in Table VII. Then, the controller developed in Section III-B obtains an optimal VM migration action, which specifies the number of type k servers to migrate from datacenter 1 to type l servers in center j , as shown in Table VIII. In addition, the controller also evaluates the result as shown in Scenario #1 of Table X. For example, a total of 419.8 kW of power is curtailed at datacenter 1, which is 119.9% of the requested. Also, a profit of more than \$442.1 can be achieved with a 95% guarantee. If the operator decides to accept PJM's signal, then datacenter 1 and the other centers involved in the VM migration plan can specify which servers to migrate out and which servers to use to accommodate the incoming VMs, respectively. The execution-time overhead of the controller was measured to be 13 secs.

C. Formulations with Risks

In the second scenario, we maintain our previous simulation parameters and use the formulation developed in Section III-C where only the bandwidth risk is considered. In Table V, the uncertainty in the available bandwidth between datacenter 1 and 7 is much higher than other links. Therefore, to avoid an unstable link and reduce the migration risk, the number of servers migrated to datacenter 7 decreases—from the previous value of 59 servers to 13 servers, as shown in Table IX. The remaining servers are mainly directed to datacenter 4 instead, because the migration times incurred on the other links have already reached the limit imposed by the 3-hour lead-time window. When the bandwidth risk is included, the new power curtailment drops a little to 405.5 kW, and the profit threshold is reduced to \$421.7, as indicated by Scenario #2 of Table X. The controller takes 20 secs to find the optimal solution.

In the final scenario, the uncertainty of LMPs, shown in Table VI, is added to the second scenario. As the controller is designed to be risk averse, it affects a tradeoff between the profit and the associated variance by reducing the power curtailed to 381.2 kW; in other words, it lowers the expected profit in order to reduce the variance band around the profit. This result is summarized by Scenario #3 in Table X. An

TABLE VIII
IN THE BASIC SCENARIO, THE NUMBER OF SERVERS MIGRATED OUT FROM DATACENTER 1 TO OTHER DATACENTERS, AND THE CORRESPONDING MIGRATION TIME ON EACH LINK.

Datcenter	2			3				4			5			6			7		
Server Type	1	2	3	1	2	3	4	2	3	4	2	3	4	1	3	4	1	3	4
1	72	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
2	0	268	0	304	0	0	25	0	203	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	55	0	59	0	0
Total Number	340			304				25			203			56			59		
Migration Time (min)	179.7			179.8				39.5			179.9			178.7			179.8		

TABLE IX
IN THE SCENARIO CONSIDERING BANDWIDTH RISK, THE NUMBER OF SERVERS MIGRATED OUT FROM DATACENTER 1 TO OTHER DATACENTERS, AND THE CORRESPONDING MIGRATION TIME ON EACH LINK.

Datcenter	2	3	4	5	6	7
Total Number	340	304	70	203	56	13
Migration Time (min)	179.7	179.8	110.6	179.9	178.7	39.6

TABLE X
SUMMARY OF TEST SCENARIOS DISCUSSED IN THE PAPER.

Scenario #	Data Center	ΔP (kW)	$\frac{\Delta P}{\Delta \bar{P}}$	Reward (\$)		Migration Cost (\$)	Diff. Elec. Cost (\$)	Distance Effect	Migration Risk	Profit Risk	Success Probability	Profit Threshold (\$)
				μ	σ							
1	1	419.8	119.9%	558.4	71.1	22.1	-22.8	19.7			95%	442.1
2	1	405.5	115.9%	539.4	68.7	21.3	-16.7	19.8	41.3		95%	421.7
3	1	381.2	108.9%	507.0	64.5	20.1	-24.0	16.9	26.7	83.3	95%	404.8
4	1	365.9	104.6%	985.2	88.8	31.1	-139.0	21.2	60.8	157.7	95%	947.1
	4	201.3	100.7%									

execution time of 6.6 seconds is used to obtain the optimal VM migration scheme.

D. Multiple Data Centers Signaled at A Time

The proposed method also works for the case where multiple datacenters receive curtailment signals at the same time. In Scenario #4 of Table X, we assume that datacenters 1 and 4 both receive signals, requiring load curtailments of 350 and 200 kW, respectively. The controller, in the same fashion, spends 54 seconds to decide an optimal VM migration policy and computes that a profit of more than \$947.1 can be achieved 95% of the time. Finally, it is possible that when n datacenters receive their own RTO signals at the same time, a solution cannot be found because the constraints in Section III are not all satisfied. In this situation, the controller would move to check if $n - 1$ of the datacenters can commit to the DR program, and so on.

V. RELATED WORK

We now discuss some recent work related to our paper in the area of load balancing on geographically distributed datacenters.

Qureshi et al. show how to exploit the volatility in the day-ahead electricity market using a set of distributed datacenters to reduce overall energy cost [20]. Datacenters in regions with high day-ahead prices are deactivated and the workload is rerouted to active ones in lower-LMP regions. The method will not work for session-based workload since it does not consider VM migration, and more importantly, a DR program

is not considered. Abbasi et al. propose a dynamic application-hosting management scheme for distributed datacenters to minimize the cumulative energy, server switching, and workload migration costs, by determining the number of active servers and the workload share for each datacenter [21]. However, the authors assume homogeneous servers with identical power consumption profiles and the applications are assumed to be hosted on a non-virtualized environment, effectively removing VM migration as a possibility. Also, instead of predicting the LMPs, they assume that these values are known a priori, and do not consider DR programs.

Irwin et al. present a staggered “blinking” approach for storage systems within a datacenter where hard disks are turned on/off so that the power consumption is adjusted based on the time-of-use LMPs or power variations from renewable energy sources [22]. Rao et al. [23] aim to minimize the total electricity cost of distributed datacenters by dispatching workload to locations with lower real-time electricity prices. However, they assume homogeneous and non-virtualized servers in each datacenter receiving the same traffic rate. Also, server switching costs (in terms of delay in turning the server on/off and transient power consumption) and network delays when rerouting the workload are not considered.

Liu et al. propose distributed algorithms to provably compute optimal routing and server provisioning decisions for Internet-scale systems [24]. Their load balancing scheme can also route the load to follow the availability of renewable sources such as wind and solar, and reduce brown energy con-

sumption. They assume that the datacenters host homogeneous servers with equal processing capacities and do not discuss server switching costs and VM migration. A DR program is considered, but is limited to only coping with dynamic electricity prices. Liu et al. evaluate the feasibility of powering internet-scale systems using mostly renewable energy [25]. They show that geographical load balancing of the incoming workload can significantly reduce the storage capacity needed to smooth the ups and downs of renewable power by routing requests to datacenters in areas that are currently generating renewable power.

Le et al. [26] propose policies to dynamically place and migrate workload across multiple geographically distributed data centers housing virtualized servers, such that the electricity costs are reduced due to the time-based differences in electricity prices. The policies model the energy costs, peak power costs, the impact of outside temperature on cooling energy consumption, and the overhead of VM migration. Chen et al. [27] propose a workload scheduling algorithm to minimize the overall brown energy consumption of geographically distributed datacenters by substituting with renewable energy, where non-interactive workloads are dynamically dispatched to VMs in different datacenters.

To summarize, the major difference between the presented framework and prior work is that, most existing techniques use some form of self scheduling in that the workload follows low LMP values (or renewable energy) in an effort to reduce the costs of operating geographically distributed datacenters, whereas we have proposed optimization frameworks that will allow datacenters to behave as CLRs within the demand dispatch paradigm for economic payoff.

VI. CONCLUSION

We have described work aimed at developing an optimization framework that allows datacenters to be treated as DR resources that can participate in wholesale energy markets and earn financial reward by reducing the consumption of electricity in response to price signals from utility companies. The risks incurred by real world uncertainties such as dynamic bandwidth and LMPs are also considered. Simulation results involving datacenters in multiple regional markets are presented to validate the framework. They show that the optimizer affects a reasonable tradeoff between the number of VMs migrated over a link and the bandwidth fluctuation on that link. Also, the migration scheme is risk-averse to the variance of electricity prices and aims to achieve a tradeoff between the expected profit and the variance associated with the profit.

ACKNOWLEDGMENT

This work was funded in part by the National Science Foundation through grant number CNS-0643888.

REFERENCES

- [1] R. Walawalkar, S. Fernandes, N. Thakur, and K. R. Chevva, "Evolution and current status of demand response (DR) in electricity markets: Insights from PJM and NYISO," *Energy*, vol. 35, no. 4, pp. 1553 – 1560, 2010.
- [2] D. Bouley, "Estimating a data center's electrical carbon footprint," April 2010.
- [3] EPA, "Report to congress on server and data center energy efficiency," July 2007.
- [4] PJM, "PJM demand side response overview," May 2012.
- [5] J. Goellner et al., "Demand dispatch intelligent demand for a more efficient grid," *Technical Report DOE/NETL- DE-FE0004001*, August 2011.
- [6] PJM, "Introduction to PJM demand response," Oct. 2008.
- [7] —, "Retail electricity consumer opportunities for demand response in PJM's wholesale markets," 2012.
- [8] ISO-New-England, "Real-time price response program," 2012.
- [9] VMware, "VMware vSphere vMotion: 5.4 times faster than Hyper-V live migration," 2011.
- [10] VMware and Cisco, "Virtual machine mobility with VMware VMotion and Cisco data center interconnect technologies," 2009.
- [11] FS, "Deploying the BIG-IP v10.2 to enable long distance live migration with VMware vSphere vMotion," 2010.
- [12] DeepStorage, "HyperIP enables VMotion over WAN," 2009.
- [13] Cisco, "Data center interconnect design guide for virtualized workload mobility with Cisco, NetApp and VMware," 2011.
- [14] —, "Data center interconnect implementation guide for virtualized workload mobility with Cisco, EMC and VMware," 2011.
- [15] D. Kusic, J. Kephart, J. Hanson, N. Kandasamy, and G. Jiang, "Power and performance management of virtualized computing environments via lookahead control," in *Proc. of IEEE Intl. Conf. on Autonomic Computing (ICAC)*, Jun. 2008, pp. 3–12.
- [16] P. Krokhmal, M. Zabarankin, and S. Uryasev, "Modeling and optimization of risk," *Surveys in Operations Research and Management Science*, vol. 16, no. 2, pp. 49 – 66, 2011.
- [17] M. Steinbach, "Markowitz revisited: Mean-variance models in financial portfolio analysis," *SIAM Review*, vol. 43, no. 1, pp. 31–85, 2001.
- [18] S. G. Makridakis, S. C. Wheelwright, and R. J. Hyndman, *Forecasting: methods and applications*, S. C. Wheelwright, Ed. Wiley, 1998.
- [19] EIA, "Electric Power Monthly," <http://www.eia.gov/electricity/monthly/index.cfm>.
- [20] A. Qureshi, "Plugging into energy market diversity," in *7th ACM Workshop on Hot Topics in Networks*, Oct. 2008.
- [21] Z. Abbasi, T. Mukherjee, G. Varsamopoulos, and S. K. S. Gupta, "DAHM: A green and dynamic web application hosting manager across geographically distributed data centers," *ACM Journal on emerging technology*, 2012.
- [22] D. Irwin, N. Sharma, and P. Shenoy, "Towards continuous policy-driven demand response in data centers," in *ACM Workshop on Green Networking*, Aug. 2011.
- [23] L. Rao, X. Liu, L. Xie, and W. Liu, "Minimizing electricity cost: Optimization of distributed internet data centers in a multi-electricity-market environment," in *IEEE INFOCOM*, 2010.
- [24] Z. Liu, M. Lin, A. Wierman, S. H. Low, and L. L. Andrew, "Greening geographical load balancing," in *ACM SIGMETRICS*, 2011.
- [25] —, "Geographical load balancing with renewables," *SIGMETRICS Perform. Eval. Rev.*, vol. 39, no. 3, pp. 62–66, Dec. 2011.
- [26] K. Le, J. Zhang, J. Meng, R. Bianchini, Y. Jaluria, and T. Nguyen, "Reducing electricity cost through virtual machine placement in high performance computing clouds," in *Intl. Conf. for High Performance Computing, Networking, Storage and Analysis*, 2011.
- [27] C. Chen, B. He, and X. Tang, "Green-aware workload scheduling in geographically distributed data centers," in *Intl. Conf. on Cloud Computing Technology and Science*, 2012.