RED-BL: Energy Solution for Loading Data Centers

Muhammad Saqib Ilyas	Saqib Raza	Chao-Chih Chen	Zartash Afzal Uzmi	Chen-Nee Chuah
SSE, LUMS	Cisco Systems	UC, Davis	SSE, LUMS	UC, Davis
saqibm@lums.edu.pk	sraza@ucdavis.edu	cchchen@ucdavis.edu	zartash@lums.edu.pk	chuah@ucdavis.edu

***Abstract*—Cloud infrastructure providers and data center operators spend a major portion of their operations budget on the electric bills. We present RED-BL (Relocate Energy Demand to Better Locations), a framework for determining an optimal mapping of workload to an existing set of data centers while considering the cost of workload relocation. Within each workload mapping interval, RED-BL solution exploits the geo diversity in electricity price markets. The temporal diversity in those markets is simultaneously exploited by considering a planning window comprising several mapping intervals.**

Using workload traces from live Internet applications and electricity prices from the US markets, RED-BL can reduce the electric bill by as much as 81% from the case when the workload is equally distributed. Compared to a single data center deployment, an average reduction of 27% in electric bill can be achieved when RED-BL uses 10 or more data centers, a common case for most operators. When compared to existing workload relocation solutions, RED-BL achieves a further reduction of 13.63%, on average. While modest, this reduction can save millions of dollars for the operators. The cost of this saving is an inexpensive computation at the start of each planning window.

I. Introduction

To enable resilient and low-latency public and private cloud services, companies such as Amazon, Google and Microsoft deploy a huge infrastructure in the form of distributed data centers. The cost of electricity needed to run these data centers accounts for a significant portion of the total capital and operation expenditure [1]. Furthermore, the fraction representing the cost of electricity is on the rise [2, 3], making it important for the cloud providers and data center operators to cut down on their electric bills.

The electricity cost for the data center network, for a given interval, depends not only on how much workload is being computed, but also on where it is being computed. Since electricity prices exhibit temporal as well as geographical diversity, it was proposed that workload could dynamically be shifted between data centers to optimize for the electricity cost. Towards that end, an operational planning window, such as a day, could be divided into multiple intervals, and a state trajectory optimization problem would be used to determine the optimal states for each interval during the planning window. Preliminary evaluations of this scheme have been reported in [4–6].

Fig. 1 shows an example of a three-interval planning cycle of mapping workload to three data centers. Only three states for each interval are shown for simplicity. Each state is represented by a rectangle, with one circle for each data center. The shading within a circle represents the amount of workload

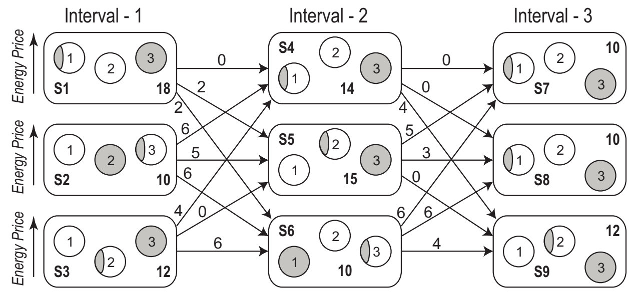

Fig. 1: A motivating example

that would be mapped to the corresponding data center if that state were chosen in that particular interval. The height of the circle within a state reflects the electricity price: a circle placed higher in the box means higher cost of electricity at that data center during that interval. The cost of operating in a particular state for an interval is also shown within the same rectangle. Using labels on the arrows between states, the figure also shows a transition cost associated with relocating workload between data centers. This *transition cost* has largely been ignored by prior work [4–6].

Previously proposed techniques for relocating workloads typically make a *greedy* choice of state in each interval [4–6]. We refer to these as Relocate Energy Demand to Cheaper Locations (RED-CL). In Fig. 1, the greedy choice corresponds to the path S2→S6→S8 with a sum of state costs equaling 30. This is clearly the lowest possible sum of state costs without considering any transition costs. However, transition costs exist in practice and may represent a significant proportion of the energy cost in a single interval. With transition costs included, the greedy solution yields a total cost of 42.

One may also consider a *static* deployment configuration where an operator selects the data centers that have the lowest average electricity price over the planning window. This corresponds to the path S1→S4→S7, with a total sum of state costs equal to 42. Since the workload mapping does not change, there are no transition costs, and hence the total solution costs is also 42. In general, with transition costs included, the static solution could be better or worse compared to the greedy solution.

The optimal solution from Fig. 1 is the path S3→S5→S9, with a total cost of 39. For this state path, the sum of state costs is 39, which is higher than the corresponding component for the greedy solution. However, the sum of transition costs is 0 resulting in an overall lower total solution cost than the static or greedy strategy. This simple example illustrates that it is important to consider the costs associated with relocating demands in operational data centers.

In this paper, we present Relocate Energy Demand to Better Locations (RED-BL), a framework for optimizing an operator's electricity costs by dynamically re-assigning workload to available data centers at discrete intervals in a planning window. This optimization considers not only the electricity cost of a particular workload assignment, but also the cost of transition from one network state to another. Note that the goal of this paper is not to accurately model the exact sources of transition costs, which can vary from provider to provider. Instead, we take a parameterized view of transition costs and study the sensitivity of workload relocation schemes to the magnitude of transition costs relative to cost of network states. Due to this abstraction, any operator should be able to apply our results to their own deployment with minimal effort.

We find that using RED-BL workload relocation solutions, an operator may save up to 81% of their electric bill, for a wide range of transition costs, compared to the case of uniformly distributing the workload among data centers. This, on average, is 13.63% better as compared to the existing RED-CL solutions. While this percent additional saving is modest, it can translate into millions of dollars of savings for large operators. To realize these savings, RED-BL requires a quick computation at the start of each planning interval. Altogether, this paper makes the following contributions:

1) RED-BL, the first electric bill minimization solution for data center operations considering the cost of workload relocation.
2) A formulation of the network state trajectory optimization problem; the solution (RED-BL) picks a sequence of network states over a *look-ahead* planning window.
3) A thorough evaluation of RED-BL and comparison with RED-CL using electricity prices from the US markets and workload data from live Internet applications, for a wide variety of operators (with number of data center varying from 1 to 33), and data centers of varying capacity. We also performed a sensitivity analysis of the RED-BL solution as the cost of activating and deactivating a data center changes.
4) To the best of our knowledge, the first study to evaluate the sensitivity of workload relocation solutions (RED-BL and RED-CL) to workload prediction accuracy, amount of over provisioning, and geographical diversity.

Our solution provides detailed operational planning information in the form of:

- A list of data centers to be kept active for each interval in the planning window, and
- The workload distribution amongst these data centers.

The rest of the paper is organized as follows. Section II presents a model for a cloud operator's network. Section III describes the experiment setup and the datasets that we used. In Section IV we present a summary of the results of our study. In Section V, we draw our conclusions.

Parameter	Description
m	Number of data centers
n	Number of intervals in a planning window
c_i	Normalized workload capacity of data center i
σ	Penalty for activating a unit capacity data center as a fraction of its energy consumption at full load in one interval
δ	Penalty for activating a unit capacity data center as a fraction of its energy consumption at full load in one interval
e_i^j	Unit cost of electricity at data center i during interval j
λ	Duration of an interval in hours
w^j	Operator's workload during interval j
x_i^j	Workload mapped to data center i during interval j
p_i^j	1 if data center i is active (either computing workload or idling) during interval j, 0 otherwise
b_i^j	1 if data center i is activated at interval j, 0 otherwise
s_i^j	1 if data center i is deactivated at interval j, 0 otherwise

TABLE I: Data Center Network Model Parameters

II. Problem Formulation

A cloud operator's infrastructure consists of several interconnected geographically diverse data centers. The workload consists of client requests for hosted applications. Every request is routed to one of the data centers in the network. For the problem of mapping the workload to data centers over several consecutive intervals in a planning window, we aim to formulate an optimization problem using variables and parameters defined in Table I.

Fan et. al. have modeled the electrical energy consumption of a data center as a linear function of the average CPU utilization of servers, with a non-zero idle power consumption [7]. Our power consumption model is equivalent but less granular: we also use normalized values for workload (0 representing no workload and 1 the peak workload) but for the entire data center rather than individual servers. Data center capacities are also expressed on the same scale. Thus, the sum of data center capacities must be at least 1 to serve the peak workload.

Let P^{idle} be the idle and P^{peak} be the peak power consumption for the entire data center network. If c_i and U_i^j, respectively, are the normalized workload capacity and the utilization of data center i, during interval j, the power consumption is $P_i^j = c_i(P^{idle} + U_i^j(P^{peak} - P^{idle}))$. The idle power consumption may be expressed as $P^{idle} = fP^{peak}$ ($0 \leq f \leq 1$). Our model thus becomes: $P_i^j = P^{peak}(f + (1-f)U_i^j)$. Dividing both sides by P^{peak}, we get the normalized power consumption: $\hat{P}_i^j = c_i(f + (1-f)U_i^j)$.

If x_i^j represents the amount of workload mapped to data center i during interval j, the utilization of data center i during interval j is given by $U_i^j = x_i^j / c_i$. Thus, $\hat{P}_i^j = c_i(f + (1-f)x_i^j/c_i)$. Multiplying this with λ and the unit price of electricity gives the electricity cost spent on computing workload at data center i during interval j. Using this model for electricity cost computation, the RED-BL optimization problem is stated as:

$$\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j (p_i^j \lambda (f + (1-f)\frac{x_i^j}{c_i}) + b_i^j \sigma + s_i^j \delta)$$

subject to:

$$x_i^j \leq c_i \quad \forall i, \forall j \tag{1}$$

$$\sum_{i=1}^{m} x_i^j = w^j \quad \forall j \tag{2}$$

$$p_i^j, b_i^j, s_i^j \in \{0, 1\} \quad \forall i, \forall j \tag{3}$$

$$p_i^j \geq x_i^j \quad \forall i, \forall j \tag{4}$$

$$b_i^j \geq p_i^j - p_i^{j-1} \quad \forall i, 2 \leq j \leq n \tag{5}$$

$$s_i^j \geq p_i^{j-1} - p_i^j \quad \forall i, 2 \leq j \leq n \tag{6}$$

$$b_i^0 = p_i^0, s_i^0 = 0 \quad \forall i \tag{7}$$

The sum of state costs (i.e., cost of computing the workload) in the optimal trajectory is represented by the first two terms in the objective function. The multiplication of the first term by p_i^j ensures that idling cost is incurred only when a data center is active. Transition costs are taken into account by considering the cost of activation and deactivation of data centers and is represented by the last two terms in the objective function.

The workload capacity constraint is given in (1). Eq. (2) ensures that all incident workload is handled, while (3) represents the binary-value constraint. Inequality (4) ensures that a data center is active whenever there is any workload mapped to it, while the rest of the constraints are related to the feasibility of activation and deactivation.

III. EXPERIMENT SETUP

In this section, we describe the experimental setup to perform a comparative study of different workload placement algorithms under various scenarios.

A. Application workload

We used an year-long trace of hourly workload for 3 social networking applications, with a subscription base of over 8 million users [8]. We sliced the trace into week-long segments and considered each slice as workload for a different application, for the same week. We normalized the sum of these trace vectors so that the peak cumulative workload corresponds to a value of 0.9. The qualitative nature of the workload is described in [9].

B. Electricity prices

We selected 33 different regions in the USA (including New York, California, Midwest, and New England) for which hourly electricity prices are available online. We used the day-ahead prices for these locations, i.e., the electricity price negotiated for the same hour on the following day. We formulated the following deployment scenarios using this dataset:

1) **Single site operator:** An operator that only has one data center at a fixed location. We considered 33 different cases in this type of deployment, corresponding to each location for which we had electricity price data.
2) **Multi-site operator:** For a multi-site operator, we considered six different deployment scenarios. Five of these scenarios represent operators with data centers at the first 10, 15, 20, 25 and 30 locations that we selected, while the sixth scenario represents an operator with a data center at all 33 locations.

Algorithm	Remarks
LI	Local optimal with idling
LD	Local optimal with deactivation
LS	Local optimal with selection
LO	Local optimal without transition costs
RED-BL	The global optimal
UNIFORM	Distribute workload equally over all data centers
STATIC_MIN	See III-C

TABLE II: Algorithms compared in our work

C. Algorithms for Workload Distribution/Relocation

In this paper, we report comparative results for seven workload placement algorithms, listed in Table II. These algorithms differ on the basis of the following criteria:

1) Data center selection: LI, LD, LS and LO pick the $\lceil w^j/c_i \rceil$ cheapest data centers for every interval to map the workload to. STATIC_MIN picks a single data center with the cheapest average electricity price over all intervals and maps workload to it during all intervals. UNIFORM distributes workload equally over all data centers, whereas RED-BL's data center selection is driven by a global optimization solver.
2) Unused data center behavior: LI keeps all data centers active during all intervals. In contrast, LD and LO deactivate any data center to which no workload is mapped during a given interval, thus avoiding idle power. With LS algorithm, a data center with no mapped load in a given interval can either be kept active (and consume idle power) or it may be deactivated, whichever option is cheaper[1]. RED-BL's choice in this regard is driven by the global optimization solver.
3) Reported total electricity cost: LO represents the electricity cost reported by the originally proposed greedy algorithm [5]. The total cost reported by this algorithm, hence, does not include transition costs. All other algorithms include transition costs, if any, in the total electricity cost.

D. Scenarios

To evaluate the above algorithms, we formulated four different scenarios. For each scenario, we ran 7 experiments (one for each day of the week) and report the average of the total electricity cost for each algorithm. Each experiment determines an operational plan for a planning window consisting of 24 one-hour intervals.

1) **Over-provisioning**: With data centers at all 33 locations, we varied c_i between 0.03 and 0.12 (in increments of 0.01). This covers a variety of operators whose workload capacity ranges from just over expected peak to almost 300% over-provisioning. We computed the total

[1]Consuming idle power may sometimes be cheaper than deactivating and subsequently reactivating a data center due to shutdown and bootup costs.

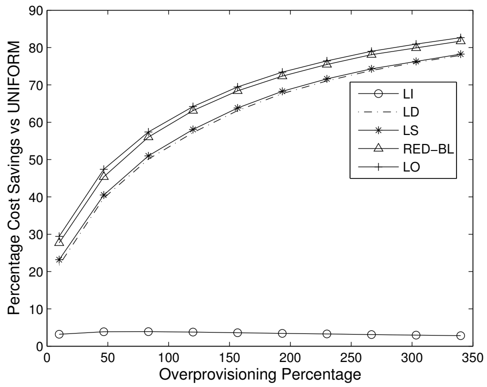

Fig. 2: Percentage savings with over-provisioning

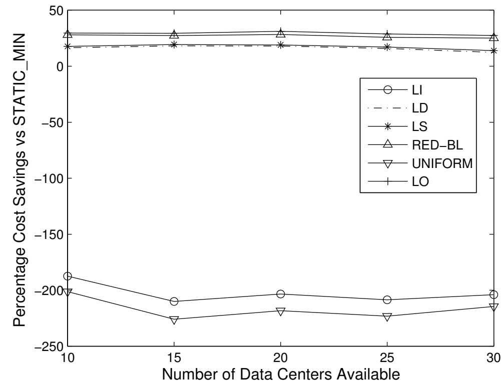

Fig. 3: Percentage savings with increased diversity

electricity cost for all algorithms, except STATIC_MIN. In this scenario, we set $f = \sigma = \delta = 0.65$.

2) **Diversity**: We compared all the algorithms as the number of data centers is varied from 10 to 30 (see 2 in Section III-B), while the total deployed capacity is fixed at 3.0 (which results in rational values for c_i in each experiment). In this scenario, we kept $f = \sigma = \delta = 0.65$.
3) **Transition cost**: We determine the total electricity cost for each algorithm (except STATIC_MIN), as the activation/deactivation overhead is varied between 0 and λ, in increments of 0.1. The lower bound on σ (and δ) implies the ideal condition of no transition overheads. We set the upper bound to λ so that the transition costs equal the cost of operating a data center at full load for an interval. A transition cost higher than this does not make sense. In this scenario, we kept $f = 0.65$.
4) **Workload estimation error**: For LS, LD and RED-BL, we added a zero mean Gaussian random variable to the workload and compared the solution cost to that obtained using error-free workload estimates. The variance of the random variable ranged from 0 (perfect estimates) to 0.12 (representing a maximum perturbation amounting to 40% of peak workload with 99.7% probability). In this scenario, we kept $f = \sigma = \delta = 0.65$. We repeated each experiment for 20 different realizations of the Gaussian error, and computed the average absolute error in the total electricity cost.

IV. Results

A. Scenario 1 (Extent of over provisioning)

The percentage savings in total electricity cost by various algorithms compared to UNIFORM are plotted against the data center capacity over-provisioning in Fig. 2. An interesting observation is that for the wide range of capacity over-provisioning that we considered, LI is able to do only slightly better than UNIFORM. This is because LI is unable to deactivate data centers that are not computing any workload, which results in significant idling costs.

The cost of the greedy strategy can be significantly improved, compared to that of LI, by deactivating unused data centers. Prior work has reported this (shown as LO) while ignoring transition costs. Such costs, however, exist and the curves for LD and LS indicate that the total cost (including transitions) is higher. For example, LS is 18.48% higher than LO, on average. The cost of RED-BL solution, however, is not too far from the LO prediction.

The reason for greater savings with RED-BL compared to LS and LD is that the latter two make a greater number of state transitions in a planning window than RED-BL. RED-BL often incurs more cost on computing workload at locations that are not the cheapest in the given interval, but compensates for this additional expense by a reduction in the number of transitions.

B. Scenario 2 (Geographical Diversity)

Fig. 3 shows the savings in electricity cost achievable using various algorithms compared to the electricity cost of STATIC_MIN against the number of data centers. Since the total deployed capacity is held fixed, per data center capacity (c_i) drops with an increase in the number of data centers.

LI and UNIFORM both fare poorly compared to the strategy of using just one data center throughout a given day. However, if the greedy solutions are allowed to deactivate data centers that are not computing any workload in a given interval, the savings due to diversity start to show (see the plots for LS and LD).

The savings achievable using LD, LS and RED-BL exhibit a slightly decreasing trend with increase in diversity. This is due to a corresponding increase in the electricity cost of computing workload. As diversity increases, the total deployed capacity being fixed, the individual data center's capacity decreases. As a result, a greater number of data centers, which on average are relatively more expensive, must be chosen by these algorithms to serve a given workload.

C. Scenario 3 (Activation/Deactivation overhead)

In Fig. 4, we observe that the electricity cost for the data center network, when using UNIFORM and LI is independent of transition costs. This is because these algorithms do not change the data center states and keep all data centers active all the time.

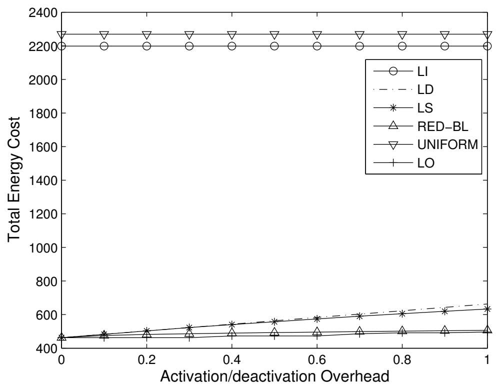

Fig. 4: Total cost vs transition overhead

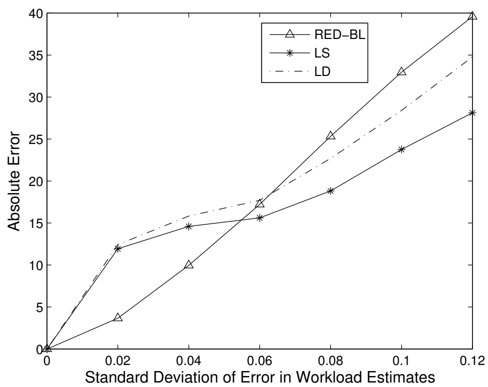

Fig. 5: Cost estimation error due to workload estimation error

On the other hand, the LS and LD adaptations of the LI algorithm scale linearly to the magnitude of transition costs. Both LS and LD also bring an average reduction in electricity cost by a factor of 4, compared to that of LI. RED-BL not only scales better than LS and LD but also achieves an electricity cost that is fairly close (only 2.95% higher, on average) to that of LO which neglects the transition costs altogether.

D. Scenario 4 (Workload estimation error)

The average absolute error in total electricity cost due to erroneous workload estimates is plotted in Fig. 5 for LS, LD and RED-BL. According to the properties of the Gaussian distribution (used as additive noise to simulate erroneous workload estimates), the swing in standard deviation from 0 to 0.12 represents a perturbation in the range from 0% to 40% of peak workload.

RED-BL exhibits a sensitivity behavior that is different from that of greedy algorithms. RED-BL's curve is almost linear. Furthermore, if the absolute perturbation in workload estimate is less than around 20% of the peak workload, the absolute error in estimated electricity costs is lower for RED-BL than that for the greedy algorithms. For worse workload estimates, however, RED-BL's estimated total cost is in greater percentage error than that for LD and LS. We do not expect that an operator's workload estimation mechanism would be erroneous by more than about 10% over a planning window of 24 hours. In that regime, RED-BL is less sensitive to such errors compared to the greedy algorithms.

V. Conclusion

Representing the mapping of workload to a set of data centers as a system state, we model a look-ahead optimization problem of finding a state trajectory that minimizes the total electricity cost of operation over a multi-interval planning window. Previous approaches to this problem ignore costs of transitions between system states.

Using live Internet application traces and electricity prices for 33 different locations in the US, we determine that transition costs can be a significant contributor to the total cost of the solution. Using RED-BL, the average electricity cost is reduced by 13.63% compared to an improved version of the previously proposed greedy solution. From the scenarios that we tested, we conclude that having geographic diversity in the data center network can reduce the average electricity cost by 27% when using RED-BL. Furthermore, RED-BL performs much better (more than a factor of 5 savings in electricity cost) than distributing workload equally over all data centers.

With the price of electricity rising much faster than the drop in server computing power per watt, the race against rising electricity costs for data centers is a tough battle. Such costs can be reduced significantly by employing RED-BL as a complementary approach to the use of energy efficient data center architectures and energy efficient server hardware.

References

[1] Albert Greenberg, James Hamilton, David A. Maltz, and Parveen Patel, "The cost of a cloud: Research problems in data center networks," in *Computer Communications Review*, January 2009, vol. 39.

[2] Kenneth G. Brill, "The invisible crisis in the data center: The economic meltdown of moore's law," Tech. Rep., The Uptime Institute, 2007.

[3] Christian L. Belady, "In the data center, power and cooling costs more than the it equipment it supports," *Electronics Cooling*, vol. 23, no. 1, Jan. 2007.

[4] Asfandyar Qureshi, "Plugging Into Energy Market Diversity," in *7th ACM Workshop on Hot Topics in Networks (HotNets)*, Calgary, Canada, October 2008.

[5] Asfandyar Qureshi, Rick Weber, Hari Balakrishnan, John Guttag, and Bruce Maggs, "Cutting the Electric Bill for Internet-Scale Systems," in *ACM SIGCOMM*, Barcelona, Spain, August 2009.

[6] Lei Rao, Xue Liu, Le Xie, and Wenyu Liu, "Minimizing Electricity Cost: Optimization of Distributed Internet Data Centers in a Multi-Electricity-Market Environment," in *IEEE Conference on Computer Communications 2010 (INFOCOM'2010)*, San Diego, USA, March 2010.

[7] Xiaobo Fan, Wolf-Dietrich Weber, and Luiz A. Barroso, "Power provisioning for a warehouse-sized computer," in *ISCA '07: Proceedings of the 34th annual international symposium on Computer architecture*, New York, NY, USA, 2007, pp. 13–23, ACM Press.

[8] Atif Nazir, Saqib Raza, and Chen-Nee Chuah, "Unveiling facebook: a measurement study of social network based applications," in *Proceedings of the 8th ACM SIGCOMM conference on Internet measurement*, New York, NY, USA, 2008, IMC '08, pp. 43–56, ACM.

[9] Muhammad Saqib Ilyas, Saqib Raza, Chao-Chih Chen, Zartash Afzal Uzmi, and Chen-Nee Chuah, "RED-BL: Relocating Energy Demands to Better Locations," 2011, Technical Report, http://suraj.lums.edu.pk/zartash/rb-tr.pdf.