Reducing Power Use and Enabling Oversubscription in Multi-Tenant Data Centers Using Local Price

Sulav Malla and Ken Christensen
Department of Computer Science and Engineering
University of South Florida, Tampa, Florida USA
sulavmalla@mail.usf.edu, christen@cse.usf.edu

***Abstract*—Multi-Tenant Data Center (MTDC) operators can increase their profits by oversubscribing their power infrastructure. However, without an appropriate power control mechanism, oversubscription can lead to expensive unplanned service outages. We propose a new power control mechanism, LOCAP (LOCAl Price for power), as a means of 1) incentivizing tenants to reduce power use to save energy and 2) controlling tenant power use to prevent over capacity cases. LOCAP is a real-time power pricing mechanism for setting local power price based on an optimization approach where the objective is to maximize aggregate tenant revenue while ensuring that total data center power consumption does not exceed capacity. Local price is set by the data center operator to reflect the current total power demand. Individual tenants use the current local price and their Service Level Agreement (SLA) parameters to determine how much power to consume. We evaluate LOCAP using workloads from real data center traces and compare with a no oversubscription case and with COOP, an existing market-based mechanism. Simulation results show that our new mechanism benefits both the tenant, by decreasing leasing costs, and the operator, by decreasing capital expense, and achieves the goal of keeping total power consumption under the data center power capacity limit while reducing overall energy use.**

I. INTRODUCTION

Data centers are expensive to build. Infrastructure costs vary according to the size, location, redundancy, and tier of the data center but generally scale linearly to the critical power [2]. Critical power is, on average, less than 60% of the total power drawn by the data center [13]. The Capital Expense (CapEx) to build a data center can range from \$11 to \$25 per watt [14]. Given the high cost of building a data center, we would like to operate it at maximum power capacity for as much time as possible. A conservative approach is to install a maximum number of servers such that the sum of the rated maximum power consumption of the individual servers is less than the power capacity of the data center. However, the actual power draw of a server is dependent on its workload with peak power consumption occurring during 100% utilization. Such high utilization occurs only rarely. The average server utilization, even in very large hyperscale data centers may be no more than 45% [13]. Furthermore, even though individual servers or racks may reach peak power capacity infrequently, it is extremely rare for the data center as a whole to reach its peak power. This is due to workload multiplexing. A study of a real data center cluster at Google with 5000 servers demonstrates that the actual power use effectively never exceeds 72% of the maximum capacity [6].

Data center operators are incentivized to oversubscribe their existing power infrastructure to increase their power utilization and to lower the Total Cost of Ownership (TCO). Power oversubscription means that the sum of peak power consumption of individual servers is greater than the power capacity limit. Power oversubscription is profitable and increases data center efficiency, but there is a risk involved of simultaneous peaking of server workloads such that power consumption exceeds the capacity. To tackle the problem of power overload, various power capping techniques have been proposed [3], [15]. Such techniques can be applied to owner-operated data centers. We need an indirect mechanism to implement this in a MTDC where the data center operator can neither know the workload nor control individual server power consumption [8].

An MTDC is a data center where the operator owns the infrastructure (building, power, and cooling) and leases the facility to multiple tenants. Tenants, who pay a monthly lease bill, may have their own servers installed and in turn provide a service to their own customer. During power overload, there is no mechanism in place for an MTDC operator to communicate this emergency to the tenants. Moreover, the operator does not have direct control over a tenant's servers or workload. Hence, oversubscribing power infrastructure of an MTDC requires a mechanism to coordinate between the operator and tenants. We propose a new control mechanism called LOCAP that enables power capping at the data center level to avoid power overload situation. LOCAP incentivizes tenants to reduce power use as well as enables power oversubscription of an MTDC. In this paper, we make two significant contributions:

- **New pricing scheme for MTDC:** In contrast to the flat monthly subscription fee charged by an MTDC operator, that is prevalent today [7], we propose to charge tenants for power use separately using a local price. Tenants pay only for energy they consume and this incentivizes tenants to use the available power wisely, thus preventing them from being wasteful.
- **Algorithm to update local price:** We model a tenant's revenue using an appropriate utility function and formulate dynamic pricing as an optimization problem of maximizing aggregate utility of tenants. We develop an iterative real-time local price update algorithm that ensures that the total power consumption of an MTDC is below the data center power limit.

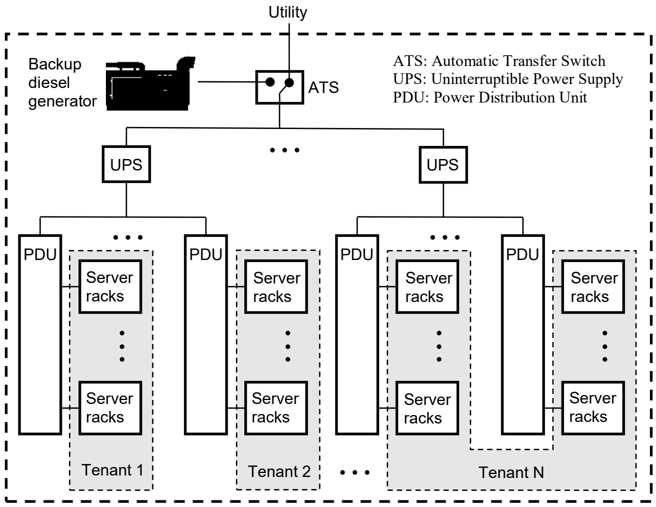

Figure 1: Power delivery infrastructure of an MTDC with *N* tenants

II. Overview

In 2014, data centers in the U.S. consumed 70 billion kWh of electricity, which is about 1.8% of the total U.S. electricity consumption [13]. Around one fourth of total data center energy consumption is from MTDCs. Thus, the MTDC is an important type of data center. Fig. 1 shows a simple MTDC power infrastructure hierarchy without any redundancy.

A. Utility Pricing

Utilities generally charge industrial customers with peak pricing comprising of an energy charge and a demand charge [7]. The energy charge is the cost of actual energy consumed by the data center (in $ per kWh) during the billing period (typically, a month). Demand charge is for the maximum average power draw for a time interval (usually 15 minutes) by the data center (in $ per kW) during the entire billing period. The power limit imposed by power utilities on large industrial customers like an MTDC is contractual. Utilities may also impose additional penalties if the actual demand exceeds the contracted demand.

B. Operator Leasing Scheme

MTDC operators generally charge tenants based upon their power subscription ($/kW/month) [7]. We will argue that such a flat pricing scheme does not encourage tenants to adopt power management techniques. Tenants may not have to keep all of their servers fully powered-up all the time, during periods of low workload they may be able to power-down some of their servers to a sleep state or operate servers at a lower frequency.

C. Tenant Service Level Agreement

Tenants, who are the service providers, have a contract in the form of a Service Level Agreement (SLA) with their customers. The SLA defines some Quality of Service (QoS) metrics that the tenant is obligated to provide to its customers throughout the period of the contract. For example, having a 95-percentile response time of 500 milliseconds [7]. The SLA typically also contains some penalty statement if the terms are violated. There is a trade-off in the costs tenants pay for power versus penalties paid to customers for failing to meet the SLA.

III. System Model

Consider an MTDC with N tenants given by the set $\mathcal{N} = \{1, 2, \cdots, N\}$. The time under consideration is divided into T time slots given by the set $\mathcal{T} = \{1, 2, \cdots, T\}$. Let the critical power capacity of the data center be C. Each tenant, $i \in \mathcal{N}$, subscribes to, and can draw, a maximum power of C_i. Since the operator oversubscribes the data center, we have

$$\textstyle\sum_{i \in \mathcal{N}} C_i \geq C. \quad (1)$$

A. Tenant and Operator Power Consumption

Each tenant i has M_i homogeneous servers out of which m_i^t servers are kept powered-on during time interval $t \in \mathcal{T}$. The rest of the servers are in sleep mode for energy conservation. A server consumes P_S power when in sleep mode, P_I power when idle, and P_M power when fully utilized. The workload of a tenant is defined for a time interval by the average request rate λ_i^t. Let the average service rate of a server be μ_i, then the average utilization of a server will be $\frac{\lambda_i^t}{m_i^t \mu_i}$ and using the power model in [6], we have

$$\begin{aligned} x_i^t &= [M_i - m_i^t]P_S + m_i^t\left[P_I + (P_M - P_I)\left(\frac{\lambda_i^t}{m_i^t \mu_i}\right)\right] && \text{if } \frac{\lambda_i^t}{m_i^t \mu_i} < 1, \\ &= [M_i - m_i^t]P_S + m_i^t P_M && \text{otherwise,} \end{aligned} \quad (2)$$

where x_i^t is the total power consumed by tenant i in time interval t. Our objective is to have $\sum_{i \in \mathcal{N}} x_i^t \leq C, \forall t \in \mathcal{T}$, that is, the sum of power consumed by all tenants to be less than or equal to the data center capacity. An MTDC operator provides cooling of server rooms along with power to tenants. The power used by the cooling infrastructure of a data center can be significant and a popular metric used to capture this is Power Usage Effectiveness (PUE). PUE is defined as the ratio of total power consumed by the MTDC to the power consumed by the IT equipment and ranges from 1.1 to 2.0 [2], [11]. The total power drawn by the MTDC from the power utility is the product of total power consumed by the IT equipment and its PUE.

B. Tenant Delay Model

In this paper, we consider tenants who provide request-response delay sensitive web services to their customers. Stated in their SLA is to have 95 percentile delay below a threshold, d_i^{th} during each time interval. We model each tenant as an M/M/n queue as in previous research [16]. The probability of sojourn time being greater than the threshold is well known and can be found in [1]. Hence, we can solve for the number of servers, required to keep this probability less than 0.05 (for 95 percentile delay).

C. Local Power Price

In this paper, we propose a new pricing structure for MTDCs to encourage tenants to be energy efficient. A flat price ($/kW/month) for infrastructure (for space and cooling) is charged as subscription fee and a real-time local power price ($/kWh), l_p^t, is charged as power consumption fee such that $L_p^{min} \leq l_p^t \leq L_p^{max}$. Here, L_p^{min} and L_p^{max} denote the minimum and the maximum value for local price. This local power price is updated every time interval and can be viewed as a power scarcity index. We assume that the tenants know the range

$[L_p^{min}, L_p^{max}]$. The operator keeps the price at a minimum during normal operation. However, in periods of power overload the local price would be increased to force tenants to reduce power use until the power overload subsides. This pricing scheme opens up the opportunity for oversubscribing the data center by varying the local price. Dynamic pricing is a well-known technique for demand side management in the utility grid [12].

D. Tenant Revenue, Cost, and Profit

The power consumption of a tenant depends on the number of powered-on servers and the workload. The number of servers kept powered-on will depend on the incoming request rate and the SLA with customers as well as the local power price and valuation of workload by the tenant. We model revenue of tenants and their valuation of customer request rate using a utility function. The utility function is denoted as $U(x, \omega)$, where x is power consumed and ω is a time varying parameter to reflect changing economic opportunity of tenant that depends on workload. The utility function represents the revenue a tenant generates by consuming x units of power. This information is private to the tenant. We consider a quadratic utility function as used in [12]

$$U(x,\omega) = \begin{cases} \omega x - \dfrac{\alpha}{2}x^2 & \text{if} \quad 0 < x < \dfrac{\omega}{\alpha}, \\ \dfrac{\omega^2}{2\alpha} & \text{if} \quad x \geq \dfrac{\omega}{\alpha}, \end{cases} \tag{3}$$

where α is a utility function parameter fixed for a tenant. As proposed earlier, tenants pay a flat monthly subscription fee, F, and a power consumption fee. The cost of consuming x units of power when the local price is l_p is $l_p x$. Hence, for a time interval the total profit of a tenant is given as

$$P(x,\omega) = U(x,\omega) - l_p x - F \tag{4}$$

where $P(x, \omega)$ is the tenant's profit and each term in the equation is in units of \$/hr. Fig. 2(a) shows utility, cost, and profit of a tenant as a function of power consumption.

Tenants would like to maximize their profit. The power consumption level that leads to maximum profit can be found by taking the derivative of (4) and setting it to zero. This gives us

$$\frac{\partial U(x,\omega)}{\partial x} = l_p = \omega - \alpha x \tag{5}$$

where $\frac{\partial U(x,\omega)}{\partial x}$ is also known as the marginal benefit [12]. From (5) tenants can determine their optimal power consumption to maximize profit for a given local price set by the operator. This optimal power consumption level is shown in Fig. 2(b), which is also known as the demand function in microeconomics. We can see that as local price increases, optimal power demand decreases. Our choice of the utility function (3) led to such a linearly decreasing demand function, which is convenient [12].

IV. Local Price Update Algorithm

It is in the best interest of all tenants as well as the data center operator, to maximize the aggregate utility of tenants. But there is a constraint that total power consumption should not exceed capacity. This problem can be formulated as the following constrained optimization problem,

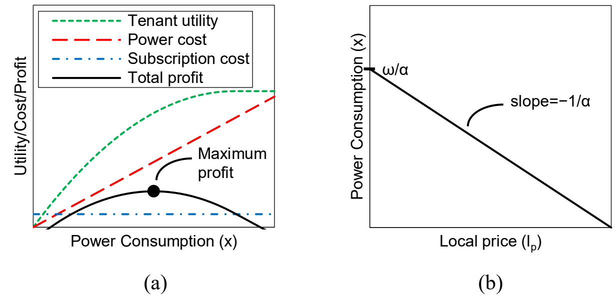

Figure 2: (a) Utility, cost and profit of tenants. (b) Demand function.

$$\max_{x_i^t \in I_i, i\in\mathcal{N}, t\in\mathcal{T}} \sum_{t\in\mathcal{T}} \sum_{i\in\mathcal{N}} U(x_i^t, \omega_i^t), \qquad \text{s.t.} \sum_{i\in\mathcal{N}} x_i^t \leq C, \tag{6}$$

$\forall t \in \mathcal{T}$, where $I_i = [0, C_i]$ is the range of possible power consumption levels by tenant i. Equation (6) is separable in t and it can be solved independently for each time interval [12]. Hence, for each time interval $t \in \mathcal{T}$, we can solve the following,

$$\max_{x_i \in I_i, i\in\mathcal{N}} \sum_{i\in\mathcal{N}} U(x_i, \omega_i), \qquad \text{s.t.} \sum_{i\in\mathcal{N}} x_i \leq C. \tag{7}$$

We can see that (7) is separable in x_i, power consumption by each tenant, for some specific time interval but is coupled by the constraint that sum of power consumption by all tenants must not exceed the power capacity. Since our objective function is concave, (7) can be solved using convex programming techniques [5] in a central fashion. A centralized solution requires the operator to know the utility function parameters of each tenant. As this information is private to each tenant, such a solution is not feasible. Hence, we need a distributed approach.

The Lagrangian [5] of problem (7) is

$$L(x,\beta) = \sum_{i\in\mathcal{N}} U(x_i,\omega_i) - \beta\left(\sum_{i\in\mathcal{N}} x_i - C\right) \tag{8}$$

where β is the Lagrange multiplier for the time interval under consideration. We can see that (8) is separable in x_i. Hence, the Lagrange dual function [5] is

$$\begin{aligned} D(\beta) &= \max_{x_i\in I_i, i\in\mathcal{N}} L(x,\beta) \\ &= \sum_{i\in\mathcal{N}} B_i(\beta) + \beta C \end{aligned} \tag{9}$$

where

$$B_i(\beta) = \max_{x_i \in I_i} U(x_i,\omega_i) - \beta x_i \tag{10}$$

and the dual problem is

$$\min \; D(\beta), \qquad \text{s.t.} \;\; \beta \geq 0. \tag{11}$$

We can observe that the dual objective function $D(\beta)$ in (9) can be decomposed into N sub-problems given in (10). These can be solved by tenants individually in a distributed fashion.

Strong duality holds, and solving the dual problem in (11) is equivalent to solving the original optimization problem in (7) [12]. The solution to dual problem, β^*, can be found and tenants can in turn find their optimal power consumption levels x_i^* by solving (10). A closer look at the objective function in (10) and

Algorithm 1 : Operator price update algorithm.

For each time interval $t \in \mathcal{T}$, at the end

Monitor power consumption x_i^t of each tenant, $i \in \mathcal{N}$.

Compute new local power price

$$l_p^{t+1} = \left[l_p^t + \Upsilon \left(\sum_{i \in \mathcal{N}} x_i^t - C \right) \right]_{L_p^{min}}^{L_p^{max}}.$$

Communicate new price to all tenants.

End for

Algorithm 2 : Reaction algorithm for each tenant $i \in \mathcal{N}$.

For each time interval $t \in \mathcal{T}$, in the beginning

Receive local power price l_p^t from operator.

Update power consumption value to

$$x_i^t = \frac{\omega_i^t - l_p^t}{\alpha_i}.$$

End for

the tenant's profit function in (4) reveals their similarity. For a tenant, solving (10) is equivalent to maximizing their profit, which is what any rational tenant would do. Optimal power consumption for maximum profit can easily be found by solving (5) as $x_i^* = \frac{\omega_i - l_p}{\alpha_i}$. Hence, if the operator sets the local price as $l_p = \beta^*$ and tenants try to maximize their own profit, we are guaranteed by strong duality that total power consumption by tenants does not exceed the capacity.

The dual problem can be solved iteratively using the gradient projection method. We update the solution by adding the negative of the gradient $\nabla D(\beta)$ after each time interval [9] as follows

$$\beta^{t+1} = \left[\beta^t - \Upsilon \frac{\partial D(\beta^t)}{\partial \beta} \right]^+ \tag{12}$$

where Υ is the step size and $[.]^+ = \max\{., 0\}$. We can get the derivative from (8) and (9) as

$$\frac{\partial D(\beta)}{\partial \beta} = C - \sum_{i \in \mathcal{N}} x_i^* \tag{13}$$

where x_i^* are solution of subproblems in (10) (the optimal power consumption level of tenants). Replacing this value of the derivative in (12) and assuming the data center operator sets the local price according to the iterative solution, that is, $l_p = \beta$, we get our final price update algorithm. To prevent local price from being arbitrarily large or small, we bound it within the range $[L_p^{min}, L_p^{max}]$. Hence, we have

$$l_p^{t+1} = \left[l_p^t + \Upsilon \left(\sum_{i \in \mathcal{N}} x_i^{t*} - C \right) \right]_{L_p^{min}}^{L_p^{max}} \tag{14}$$

where $[.]_l^h = \min\{h, \max\{., l\}\}$.

Note that $\sum_{i \in \mathcal{N}} x_i^{t*}$ is the total power consumption by the data center (demand) and C is the power capacity (supply). Our price update rule and tenant's reaction, also presented in

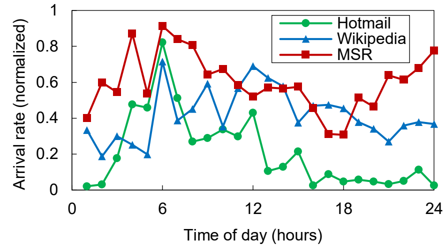

Figure 3: Workload

Algorithms 1 and 2, is intuitive in the sense that when demand exceeds supply, we increase the price and when supply exceeds demand, we decrease the price, proportional to the difference between demand and supply. The minimum value of local price, L_p^{min}, must be such that operator can transfer its electricity cost onto the tenants.

V. EVALUATION

In this section, we describe our trace-based simulation for evaluating LOCAP and comparing it with an existing method.

A. Experimental Setup

We performed a discrete-event simulation for one day with a 5-minute time interval (288 time slots). We consider an 8 MW multi-tenant data center having a PUE of 1.8 [2], [13], hence, it has a contract demand of 14.4 MW with the power utility. We consider a simplified peak pricing scheme employed by the utility with an energy charge of \$0.06 per kWh, demand charge of \$11.58 per kW, and an additional penalty charge of \$11.58 per kW for exceeding the contracted demand [10]. A 25% oversubscription has 3% probability of overloading (43.2 minutes a day) [8], but our control mechanism decreases this duration. We assume that the MTDC power infrastructure can sustain such brief power overload without service disruption. The data center operator leases out its infrastructure to tenants at a flat subscription fee of \$100/kW/month. Additionally, there is a separate local power price that the operator varies between \$0.1/kWh and \$0.5/kWh which is updated every time interval according to (14). We take the step size, Υ, to be 0.1.

We consider that there are three tenants, each with 10,000 homogeneous servers having sleep, idle and maximum power consumption as $P_S = 10$ watts, $P_I = 145$ watts, and $P_M = 330$ watts, respectively [13]. Hence, the tenants are subscribing to 3.33 MW of power and there is about 25% oversubscription. We let ω vary in the range [0.1, 10.0] and is decided by the tenants individually, proportional to their workload such that enough servers are kept powered-on to meet the SLA under minimum local price. The power consumption level, x, is calculated from the number of servers required using (2). Hence, one way of setting ω is as, $\omega = L_p^{min} + \alpha x$ and is updated every time interval. Similarly, α is in the range [0, 3] and for a similar power consumption level, a higher value implies higher revenue and less sensitive to the changes in local price. For our simulation we fix α for the three tenants to 2.4, 0.8, and 0.4, respectively.

We use workload traces from [11], shown in Fig. 3, which have hourly varying traces from Hotmail, Wikipedia, and Microsoft Research (MSR), for tenants 1, 2, and 3, respectively.

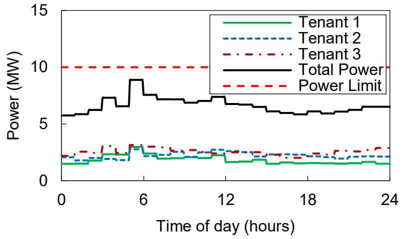

Figure 4: Power consumption for NOOV

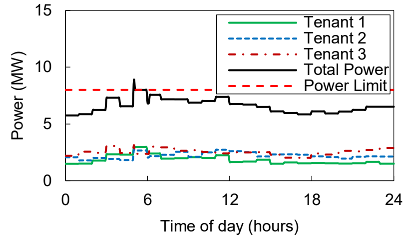

Figure 5: Power consumption for COOP

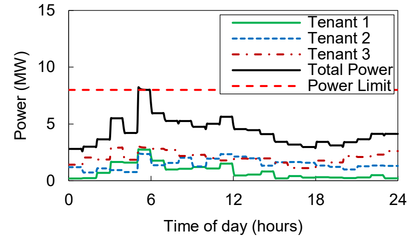

Figure 6: Power consumption for LOCAP

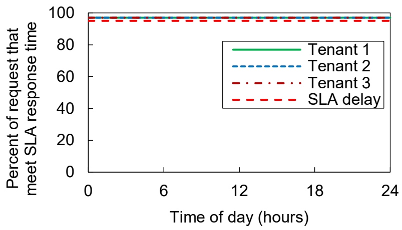

Figure 7: Delay performance for NOOV

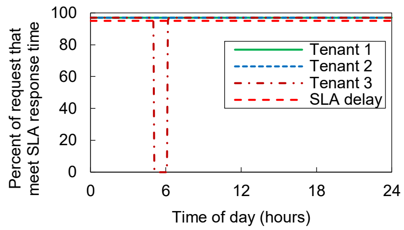

Figure 8: Delay performance for COOP

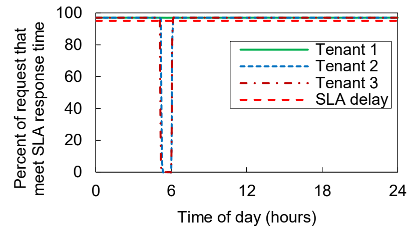

Figure 9: Delay performance for LOCAP

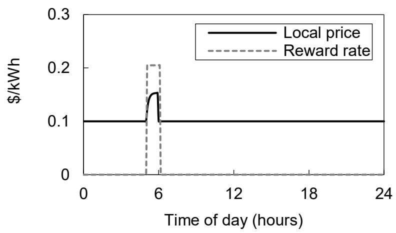

Figure 10: Real time local price for LOCAP and reward rate for COOP.

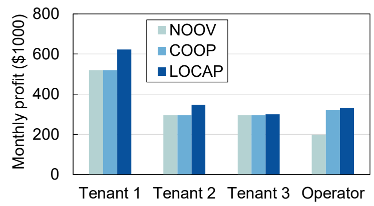

Figure 11: Monthly profits for tenants and operator under different cases.

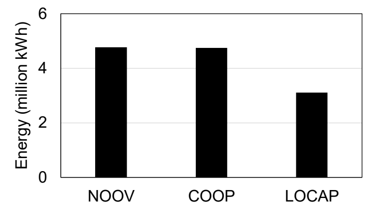

Figure 12: Monthly energy consumption under different cases.

These traces are normalized with respect to the tenant's maximum service capacity and scaled to have an average utilization of 20%, 40%, and 60%, respectively. We synthetically increased the workload at the 6th hour to create a power overload situation. We take the average service rate of a server to be $\mu = 10$ requests per second for all tenants. The tenants are providing a request-response type of web service to their customers. To avoid the request queue from being infinitely large when the system is unstable ($\rho \geq 1$), we limit the central queue size to 100 times the number of servers for each tenant. Tenants have an SLA with their customer to keep the 95 percentile delay under 350 milliseconds during each time interval. Each tenant keeps a sufficient number of servers powered-on to meet SLA. Violation of the SLA by a tenant will result in a penalty paid back to their customer denoted by decreasing revenue as per the utility function. We consider the following two baselines to compare LOCAP against.

NOOV (NO OVersubscription): This is the case when there is no power oversubscription of the data center and no control mechanism in place. For this case, we have a 10 MW data center instead of 8 MW, which significantly increases the CapEx.

COOP (CO-Ordinated Power management): This is a market mechanism approach that reward tenants for power reduction during power overload situation as described in [8]. This can be viewed as the operator buying back power from the tenants during power overload. We set maximum reward rate to \$0.5/kW/hr which is consistent with the range chosen for local price in the case of LOCAP.

In both the cases, the operator leases space at a flat fee of \$150 per kW per month [7] and there is no incentive for the tenants to reduce power during normal operation.

B. *Experimental Results*

Fig. 4, Fig. 5, and Fig. 6 show the power consumption level of individual tenants and their total for NOOV, COOP, and LOCAP, respectively. A power overload occurs at 6th hour of the simulation, which is when both mechanisms kick in to prevent sustained power overload. COOP starts by offering a reward rate while LOCAP reacts by increasing the local power price. The reward rate and the local price can be seen in Fig. 10. These results show that COOP and LOCAP both cap total power consumption and enable power oversubscription of an MTDC. However, there is one fundamental difference. In LOCAP, since we charge tenants separately for power, it encourages tenants to put unnecessary servers to sleep even during normal operating conditions. Whereas, in the case of COOP, the incentive for power reduction exists only during power overload situations during which the operator actively offers a reward rate. This distinction can be observed between Fig. 5 and Fig. 6. The total power consumption in case of LOCAP is always less than COOP at each time interval. Fig. 12 compares monthly energy consumption for the three cases. We can see that LOCAP reduces energy consumption by 34% compared to the baselines.

Fig. 7, Fig. 8, and Fig. 9 show the delay performance of the tenants under different cases. The percentage of requests that meet the delay threshold is calculated and any point above the 95 percent line denotes the SLA being satisfied. In the case of NOOV, the SLA is always met, whereas for COOP and LOCAP we have some tenants not meeting the SLA during the power overload situation. Such short-term performance degradation cannot be avoided, and might be preferable to a power failure, during power emergencies [8]. Hence, an oversubscribed MTDC is suitable for tenants who can tolerate occasional performance loss in exchange for a lower monthly cost.

Fig. 11 shows the monthly profits for the operator and three tenants under different cases. We can see that the operator profit for oversubscribed cases increases by about 64% compared to that of NOOV. Revenue and cost of tenants under NOOV and COOP were similar, leading to similar profits. However, for the case of LOCAP, while revenue was similar, monthly payment by a tenant to the operator decreased by 13%, on average, from \$150 per kW to about \$130 per kW.

VI. Related Work

Much work has been done to investigate power capping techniques which can prevent power overload situation in data centers [3], [15]. However, such work has focused on owner-operated data centers and is not suitable for MTDC. Recent studies [7], [8], [11] have focused on MTDC to have coordination between operator and tenants. These works are based on incentive mechanisms where the operator offers financial reward to tenants for power reduction. The work in [8], which introduces COOP that we have compared against, is most closely related to our work in the sense that both try to solve the same problem. We note that COOP is a bidding scheme in which tenants are rewarded for power reduction. A proper baseline is required to measure power reduction. COOP proposes current power consumption of tenants as their baseline [8]. This could encourage tenants to be wasteful, when a reward is not in effect, and set their baseline high such that when a reward is announced, they can reduce more power and gain larger rewards. The choice of baseline may be a weakness of such schemes [4].

In [17], a dynamic pricing scheme is developed to propagate operator's electricity cost onto tenants. However, they do not consider data center power capacity in their problem formulation, which should not be exceeded for safe power oversubscription. The aggregate utility maximization approach used in our work has been used for demand side management in smart grids [12] as well as for flow control in network links [9]. Our work differs from [9] and [12] in that we bound the real time local price from being arbitrarily small or large. If the local price is not bounded, a very small value would mean that the MTDC operator cannot transfer its electricity cost onto the tenants, incurring loss. Conversely, if the local price is very large, tenants may have to pay unreasonably high fees for electricity. To the best of our knowledge, LOCAP is the first dynamic local pricing control mechanism to be proposed in the literature that enables power oversubscription as well as promotes energy saving in MTDC.

VII. Conclusion

In this paper we have shown how local power price can serve to 1) enable MTDC operators to oversubscribe power and prevent over-capacity situations from occurring, and 2) reduce overall energy consumption of an MTDC. With about 16 billion kWh electricity consumed by MTDCs in the U.S. [13], a 34% energy reduction achieved by LOCAP could translate into 5.4 billion kWh of energy savings annually in the U.S. alone (if LOCAP were widely adopted).

Acknowledgment

We would like to thank Mohammad A. Islam and Shaolei Ren for providing us with the workload trace and to Bruce Nordman for his contributions on local price. We also thank the anonymous reviewers for their valuable feedback.

References

[1] I. Adan and J. Resing, *Queueing Systems*, Eindhoven University of Technology. Department of Mathematics and Computing Science, 2015.

[2] L. Barroso, J. Clidaras, and U. Hölzle, "The Datacenter as a Computer: An Introduction to the Design Of Warehouse-Scale Machines." Synthesis Lectures On Computer Architecture, Morgran & Claypool Publishers, 2013.

[3] A. Bhattacharya, D. Culler, A. Kansal, S. Govindan, and S. Sankar, "The Need for Speed and Stability in Data Center Power Capping," *Third International Green Computing Conference* (IGCC'12), 2012.

[4] S. Borenstein, M. Jaske, and A. Rosenfeld, "Dynamic Pricing, Advanced Metering, and Demand Response in Electricity Markets," University of California Energy Institurte: Center for the Study of Energy Markets, 2002.

[5] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2004.

[6] X. Fan, W.-D. Weber, and L. A. Barroso, "Power Provisioning for a Warehouse-Sized Computer," *International Symposium on Computer Architecture*, 2007.

[7] M. Islam, H. Mahmud, S. Ren, and X. Wang, "Paying to Save: Reducing Cost of Colocation Data Center via Rewards," *IEEE International Symposium on High Performance Computer Architecture*, 2015.

[8] M. Islam, X. Ren, S. Ren, A. Wierman, and X. Wang, "A Market Approach for Handling Power Emergencies in Multi-Tenant Data Center," *IEEE International Symposium on High Performance Computer Architecture*, 2016.

[9] S. Low and D. Lapsley, "Optimization Flow Control–I: Basic Algorithm and Convergence," *IEEE/ACM Transactions on Networking*, Vol. 7, No. 6, December 1999, pp. 861-874.

[10] Middle Tennessee Electric Membership Corporation, "Rate Structure For Business and Industry", 2017, http://www.mtemc.com/pages.cfm /name/b04_rates.

[11] S. Ren and M. Islam, "Colocation Demand Response: Why Do I Turn Off My Servers?," *International Conference on Autonomic Computing*, 2014.

[12] P. Samadi, A. Mohsenian-Rad, R. Schober, V. Wong, and J. Jatskevich, "Optimal Real-Time Pricing Algorithm Based On Utility Maximization For Smart Grid," *IEEE International Conference on Smart Grid Communications*, 2010.

[13] A. Shehabi, S. Smith, N. Horner, I. Azevedo, R. Brown, J. Koomey, E. Masanet, D. Sartor, M. Herrlin, and W. Lintner, "United States Data Center Energy Usage Report." Lawrence Berkeley National Laboratory, Berkeley, California, 2016.

[14] W. Turner IV and K. Brill. "Cost Model: Dollars per kW Plus Dollars per Square Foot of Computer Floor," Whitepaper, Uptime Institute, 2008.

[15] Q. Wu, Q. Deng, L. Ganesh, C. Hsu, Y. Jin, S. Kumar, B. Li, J. Meza, and Y. Song, "Dynamo: Facebook's Data Center-Wide Power Management System," *IEEE International Symposium on Computer Architecture*, 2016.

[16] L. Yu, T. Jiang, and Y. Cao, "Energy Cost Minimization for Distributed Internet Data Centers in Smart Microgrids Considering Power Outages," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 26, No. 1, January 2015, pp. 120-129.

[17] Y. Zhan, M. Ghamkhari, D. Xu, and H. Mohsenian-Rad, "Propagating Electricity Bill Onto Cloud Tenants: Using a Novel Pricing Mechanism," *IEEE Global Communications Conference (GLOBECOM)*, 2015.