

Architecting Efficient Peak Power Shaving Using Batteries in Data Centers

Baris Aksanli and Tajana Rosing
Computer Science and Engineering Department
UC San Diego
La Jolla, CA
{baksanli, tajana}@ucsd.edu

Eddie Pettis
Google Inc.
Mountain View, CA
epettis@google.com

Abstract— Peak power shaving allows data center providers to increase their computational capacity without exceeding a given power budget. Recent papers establish that machines may repurpose energy from uninterruptible power supplies (UPSs) to maintain power budgets during peak demand. Our paper demonstrates that existing studies overestimate cost savings by as much as 3.35x because they use simple battery reliability models, boolean battery discharge and neglect the design and the cost of battery system communication in the state-of-the-art distributed UPS designs. We propose an architecture where batteries provide only a fraction of the data center power, exploiting nonlinear battery capacity properties to achieve longer battery life and longer peak shaving durations. This architecture demonstrates that a centralized UPS with partial discharge sufficiently reduces the cost so that double power conversion losses are not a limiting factor, thus contradicting the recent trends in warehouse-scale distributed UPS design. Our architecture increases battery lifetime by 78%, doubles the cost savings compared to the distributed design (corresponding to \$75K/month savings for a 10MW data center) and significantly reduces the decision coordination latency by 4x relative to the state-of-the-art distributed designs.

Keywords— data center, batteries, peak power, energy cost

I. INTRODUCTION

Warehouse-scale data centers consume several megawatts and require careful power provisioning to ensure that costly power infrastructure is utilized effectively [1]. These data centers typically enter long-term power contracts and are charged market prices when exceeding their contract. The overages may be five times more expensive than their contracted rates [2]. Data centers often size their contracted power based on the expected peak power to avoid costly overages. The basic problem with power provisioning involves using as much power capacity as possible without exceeding a fixed power budget. Although individual machines may consume peak power, entire clusters of machines rarely operate at peak power simultaneously [1]. Several studies proposed peak shaving (capping) to increase power utilization [3], [4], while maintaining power budgets and amortizing capital expenditures over more machines [5].

Many mechanisms have been proposed for peak shaving, including dynamic voltage and frequency scaling (DVFS) [1], [6], virtual machine power management [7], online job migration [8], [9], [10], and batteries [3], [5], [4]. Among these, batteries are particularly useful as they do not introduce the performance overhead associated with meeting the power budget. This is especially critical during the peak

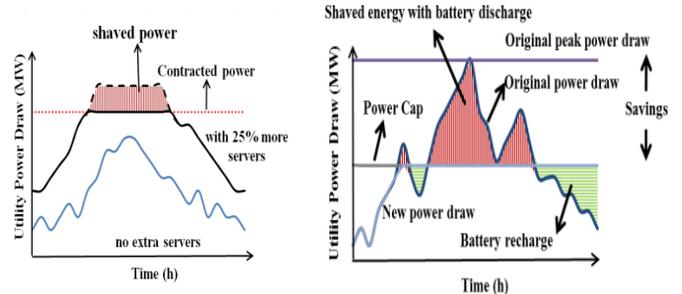


Figure 1. Sample peak shaving a) with and b) without extra servers

user demand. Battery-based peak shaving instead employs an uninterruptible power supply (UPS) to power machines.

Figure 1 illustrates two different strategies for using peak power shaving. The horizontal axes represent a 24-hour interval and the vertical axes show the aggregate power consumption. In Figure 1-a, the dotted horizontal line denotes the contracted power for the data center. The lower curve indicates the power consumption of a nominal size data center without peak shaving. A significant amount of provisioned power is wasted during low activity periods, resulting in lower profit. The upper curve adds extra servers and handles oversubscribed power with peak shaving, so that the power utilization is higher. Peak shaving prevents the power consumption from exceeding the contracted energy costs shown by the shaded region. The dashed line illustrates how much power the data center would consume without peak shaving, which would then incur as much as 5x [2] higher costs. Peak shaving increases the revenues by adding more machines to service more users and prevents utility-facing power consumption from exceeding the provisioned power with no performance cost.

Figure 1-b uses peak shaving just to decrease the level of contracted power without increasing the number of servers. The upper horizontal line represents the original peak power demand and the lower one shows the power cap. The difference between the original power draw and the power cap corresponds to energy savings as the data center can contract for less power. If the power demand is greater than the power cap, the batteries provide energy. During low power demand, the batteries recharge to regain energy in preparation for the next peak.

In state-of-the-art (SoA) work, if the data center uses a centralized UPS, the entire circuit is switched to battery until the batteries exhaust their capacity or the peak subsides. This technique is useful primarily with short pulses (a few minutes long) due to low battery capacity [2]. Recent trends in data centers focused on distributed UPS architectures,

where individual machines [11] or collection of racks [12] have their own UPS. This architecture shaves power more effectively due to the finer granularity but only works for data centers willing to implement the non-standard power architecture [5].

The main disadvantage of a centralized UPS design is the double AC-DC-AC conversion, leading up to 35% energy loss. The distributed design can avoid this double conversion by taking batteries next to the servers. Recently, DC power distribution in data centers has been proposed as a solution to decrease the conversion losses. In this paper, we also analyze the conversion losses of these different designs and quantify the effects the losses have on peak shaving capabilities.

We revisit the analyses for existing peak shaving designs using more realistic battery models and find that the benefits of peak shaving may be overestimated by up to 3.35x with simplistic models, resulting in unacceptably short peak power shaving times of only several minutes, for the centralized lead-acid UPS designs. Existing approaches discharge batteries in a “boolean” fashion: the entire data center power domain is fully disconnected from the utility power and supplied from the UPS. As a result, batteries discharge at much higher currents than rated, which lowers battery lifetime and raises the cost.

Distributed UPS design addresses this issue partly by providing the ability to discharge only a subset of batteries in a data center at a time and by using lithium iron phosphate (LFP) batteries which have both higher energy capacity and 5x more charge/discharge cycles than lead-acid (LA) batteries. The individual batteries are directly connected to servers, but still operate in boolean mode, leading to lowered battery lifetime and higher cost. Also, distributed batteries require coordination to provide the best performance. Palasamudram et al. [4] assume a centralized control mechanism and do not model the effects of coordination in their study. Kontorinis et al. [5] analyze the peak shaving performance of control mechanisms placed at different levels of power hierarchy. They conclude that the centralized controller for distributed batteries performs the best but do not comment on the feasibility of this centralized solution for a large scale system. In our estimates, the response time of a centralized controller can take up to multiple seconds which may be too long to meet the power thresholds.

A key insight that we leverage in our proposed new peak power shaving architecture is that the ideal design should have the minimum management overhead of the centralized UPS with the capability to provide “just enough” current to the data center, at a level that optimizes the individual battery lifetime. We accomplish this with a centralized UPS architecture using grid-tie inverters to partially power loads (in contrast to previous boolean discharge), so that the battery capacity decreases super-linearly with respect to discharge current [13], thus enabling the partial discharge architecture to overcome the efficiency problems associated with the state-of-the-art solutions. Our centralized grid-tie solution has 78% longer battery lifetime and doubles the cost savings compared to the best SoA distributed designs. Also, since the batteries are placed together, the communication overhead is reduced by 4x.

II. RELATED WORK

High demand and job criticality make energy a major problem for data center operators. There is a large body of work focusing on improving energy efficiency with local and global solutions. The former includes applying power shaving mechanisms such as DVFS [1], [14] and virtual machine-based power management [7] where the latter generally leverages the differences in energy prices and moves the jobs to places with cheaper energy [8], [9], [10]. But, all of these solutions adversely affect performance, e.g. DVFS slows down the applications; consolidation and migration undergo network delays.

In contrast, recent work proposes batteries to reduce the peak power of data centers with no performance overhead. The first approach is to use the existing batteries within the centralized UPS [2]. However, this method is applicable to only short peaks because the UPS powers the entire data center. In addition to batteries, Wang et al. [15] analyze flywheels and ultra-capacitors for peak shaving. Kontorinis et al. [5] and Palasamudram et al. [4] propose overprovisioned distributed batteries to sustain longer peaks. This design leads to finer grained battery output control. But, batteries require high discharge current since each one powers an entire server. High discharging current reduces both the effective battery capacity and the useful battery lifetime [16]. These publications cannot capture the negative effects of high discharge currents due to simplistic battery models and overestimate the battery lifetime. The distributed UPS implementations also do not study the overhead of management of distributed battery system at large scale.

Grid-tie inverters mainly convert DC energy generated by renewable sources into AC and feed it into the grid [17]. They allow excess DC energy to be sold back to the grid where net-metering is available. They are also used with batteries and UPS devices in grid-interactive systems for local storage and emergency response. In contrast, we propose combining battery power with the grid through grid-tie inverters during peak power periods in data centers. This achieves finer battery output control without distributing the batteries to servers, decreasing the system complexity and increasing the battery lifetime.

III. STATE-OF-THE-ART PEAK POWER SHAVING ISSUES

This section analyzes the key architectural challenges for peak shaving using batteries to have a cost and energy efficient peak shaving mechanism. We consider battery placement, power distribution type, battery chemistry, and a battery performance model. The battery placement decision and distribution system affect conversion losses. The battery chemistry dictates the peak shaving capacity. Accurate battery performance models are necessary to produce informed decisions on cost.

A. Battery Placement Designs

There are two battery placement architectures: centralized and distributed. The centralized design uses batteries within the data center-level UPS and does not require additional power equipment or infrastructure. A

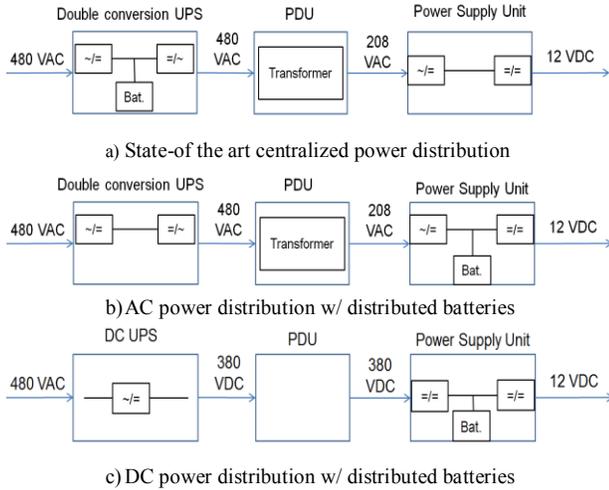


Figure 2. Different power delivery options with centralized and distributed battery placements

common power delivery hierarchy for this design using AC distribution is shown in Figure 2-a. When peak shaving occurs, the battery powers the entire data center, discharging the batteries at high rate. According to Peukert’s Law, this drains battery capacity very quickly. Also, both the AC-DC-AC double conversion in UPS and the losses on the power delivery path result in up to 35% energy loss. These losses reduce both UPS efficiency and useful battery capacity. We analyze the effects of these in more detail in section VII.

The distributed design co-locates the servers and batteries and eliminates the DC-AC battery power conversion [5], [4]. A sample design with is shown in Figure 2-b. Each server may be switched to battery independently. This leads to finer grained control of the total battery output because only a fraction of the servers are operating on battery at any given time. Together, conversion efficiency and fine-grained control permit longer peak shaving than traditional centralized designs.

In Figure 3, we compare the power shaving capabilities of the SoA centralized and distributed designs during a fixed-magnitude spike in demand without considering conversion losses. We assume each server has a 20Ah LA battery in the distributed design because that is the maximum size that can fit in a rack [5]. The centralized design has equivalent aggregate capacity to the distributed batteries. In Figure 3-a, the x-axis illustrates a range of peak server power values. We assume a provisioned power of 255W per server. This value limits the power consumption of the entire data center to $255 \times (\# \text{ servers})$. The y-axis represents the peak shaving duration corresponding to different peak power spikes. We illustrate the fixed peak power magnitude and peak power threshold in Figure 3-b. In this figure, the power curve of a data center consists of two long pulses: the peak pulse and the low pulse. The resulting power curve after peak shaving is mostly linear, having the value of the provisioned power. We define the duration batteries can sustain a specific peak pulse as the peak shaving duration. Figure 3-a has two curves showing the peak shaving durations for both centralized [2] and distributed [4] designs with different peak pulses. The former cannot scale its peak

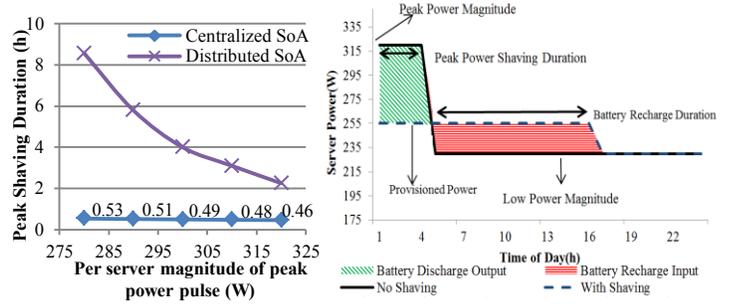


Figure 3. a) Peak shaving capabilities of different designs b) Illustration of fixed peak magnitude and peak shaving duration

shaving duration for lower magnitude peaks, whereas the latter can throttle the battery energy. The latter reduces peak power even for higher peak spikes, outperforming the centralized design by 5x when shaving 25% above provisioned power.

The success of the distributed design is due to its finer grained battery power control, but each battery still needs to power the entire server. High current reduces the effective battery capacity and reduces battery lifetime, increasing the cost. The existing distributed architectures do not account for these negative effects. In fact, our work shows that the average battery lifetime of the distributed design can be overestimated by up to 2.44x when batteries are not modeled accurately. If the battery discharge current could be shared among a group of batteries, the negative effects of high individual discharge currents would be reduced. We discuss our architecture that supports this capability in Section IV.

TABLE I. GROUP SIZES, EQUIVALENT HIERARCHY LEVEL AND THE BEST PEAK SHAVING PERFORMANCE FOR EACH GROUP [5].

Hierarchy Level	Size of a group	Best Peak Shaving
Server	1	10%
Rack	20	12%
PDU	200	16%
Cluster	1000	19%

Previous studies on distributed batteries [5], [4] assume a centralized control mechanism to obtain the best peak shaving performance with them. Palasamudram et al. [4] do not actually model a controller but their solution depends on the coordination among all the batteries, implying centralized control. Kontorinis et al. use controllers deployed at different levels of power hierarchy. Table I shows the different hierarchy levels used in that study and the corresponding number of batteries each controller needs to manage. Table I also shows the best peak shaving percentages obtained with each level of controller. Kontorinis et al. conclude that a centralized control mechanism is required to get the best performance of the distributed batteries. But, since the batteries are distributed to the servers, the centralized control mechanism needs to use the data center interconnect to manage the batteries. Kontorinis et al. do not analyze the effects of data center interconnect delays.

Currently, data centers distribute AC power because it is easy to deliver and transform. This requires multiple conversions in the power delivery hierarchy (Figures 2-a, b), such as AC-DC-AC conversions in a centralized UPS and

AC-DC conversion in the server power supply. These conversions reduce the efficiency of the centralized battery output and the distributed battery input. The former reduces the useful discharge time of the battery, and the latter leads to longer recharges.

In contrast, DC power distribution has been proposed to improve energy efficiency [18], [19]. The AC utility input is converted to DC once within a centralized DC UPS. Delivery and transformation are handled using DC. The DC option aids UPS-based peak shaving because it eliminates multiple AC-DC conversions, and up to 35% energy loss on the power delivery path. Figure 2-c shows a sample DC power distribution system with distributed batteries. This design reduces power distribution losses by up to 50% compared to the AC distribution (section VII.B). Despite its advantages, DC is not common, as it requires a new power infrastructure. It is a good option for new data centers but impractical for existing ones as the entire power distribution system must be redesigned.

B. Battery Model and Chemistry Selection

Peak shaving using batteries needs accurate estimates of battery's physical behavior. This section demonstrates how we calculate the useful battery capacity over time and estimate its depth-of-discharge (DoD) along with its available capacity after recharging and discharging. The available battery capacity at a given time is defined as the state-of-charge (SoC) and reported as a percentage of the maximum capacity. State-of-Health (SoH) quantifies the maximum deliverable capacity of a battery over time as a percentage of its initial capacity.

There are several studies estimating battery SoC and SoH, especially for mobile devices, e.g. [20] [21]. In our work, we combine a few models to both estimate the physical properties of the batteries and capture the negative effects of high discharging currents. Coulomb counting method presented in [22] describes the relation between DoD level and SoH. We take the model described in [16] to capture the effects of high discharge currents on SoH. We also include Peukert's law which states that the effective capacity of a battery decays exponentially with increasing discharging current [13]. The main benefit of this model is its simplicity and ability to easily leverage it in a large scale installation as it requires only voltage and current readings for all the calculations. We start describing our model by first calculating released capacity during a discharge event:

$$C_{released} = \Delta t * I_{discharge} \quad (1)$$

where Δt is the length of the time interval and $I_{discharge}$ is the discharge current. DoD is computed as:

$$DoD = \frac{C_{released}}{C_{eff}} \quad (2)$$

where C_{eff} is the effective capacity:

$$C_{eff} = C_R * \left(\frac{C_R}{I_{discharge} * H} \right)^{k-1} * \frac{SoH}{100} \quad (3)$$

where $I_{discharge}$ is the discharging current and C_R is the rated capacity. We use H to denote rated discharge time in

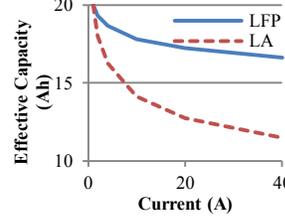


Figure 4. Effective capacity of 20Ah LA & LFP batteries

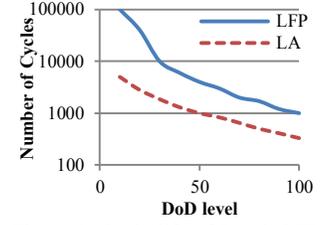


Figure 5. Cycle life of LA & LFP batteries rated at 20h [43], [44]

terms of hours and obtain its value from the data sheets, which is generally 20 hours [13]. Peukert's exponent is shown by k , which changes depending on the battery type. For LA batteries, the typical value is around 1.15 whereas for LFP batteries it is 1.05 [23]. Effective capacity is also scaled with SoH value to reflect the capacity loss as the battery is used. The DoD is subtracted from the SoC at the end of each interval. When discharging ends, we save the total DoD value during that discharge period, DoD_{final} as $(100 - SoC)\%$.

Peukert's law states the effective capacity of a battery decreases with higher discharge current. Figure 4 shows this negative effect on 20Ah LA and LFP batteries. The horizontal and vertical axes show the effective battery capacity and discharging current respectively. The effective capacity of the LA battery decreases faster due to its greater nonlinear behavior, represented by a larger Peukert exponent. At 40A, corresponding to 2C rate for both of these batteries, the LA battery loses 42% of its nominal capacity, but the LFP battery loses only 15%.

We update the battery SoH after a complete charge/discharge cycle [22]. This update depends on the battery chemistry, determining Peukert's exponent, C_{eff} and DoD_{final} . The number of charge/discharge cycles decreases with deeper discharges, represented by a larger DoD_{final} value. We use a lookup table derived from effective capacity graphs provided in commonly available battery data sheets; similar to Figure 5 for each battery chemistry to define the effects of DoD_{final} .

In Figure 5, the horizontal axis shows the DoD level for charge/discharge at 20h discharge rate, which is defined as the current that drains the battery in 20h. The vertical axis is on a log scale and illustrates the number of cycles a battery can provide for a particular DoD level. As the battery is discharged deeper in each cycle, the available number of charge/discharge cycles decreases exponentially. LFP batteries provide 5x more cycle life compared to LA batteries in average.

We normalize the effect of one cycle with DoD_{final} value to calculate its impact on the battery lifetime. The battery lifetime is defined as the interval in which battery SoH is greater than a state of health value which determines when the battery is dead, SoH_{dead} . Battery manufacturers generally recommend 80% for this value [24] [25], i.e. the battery is considered dead if the maximum capacity it can provide falls below 80% of its rated capacity. If the battery has $Cycles_{DoD_{final}}$ cycles with DoD_{final} value, the battery SoH is updated as:

$$SoH = SoH - (100 - SoH_{dead}) * \frac{1}{Cycles_{DoD_final}} * \frac{C_R}{C_{eff}} \quad (4)$$

This equation normalizes the effect of one cycle with DoD_{final} over the battery lifetime and penalizes high discharge currents.

Batteries generally include a battery management unit that both manages and monitors the voltage and the current of the battery. This unit makes it practical to use our model as it requires only voltage and current measurements of the battery. In contrast, the simple battery model used by previous work [3], [4] does not calculate C_{eff} . They use nominal battery capacity, C_R , to compute DoD. This leads to up to 42% overestimated discharge duration and thus, overestimated peak shaving capabilities. They also assume the same battery capacity over its lifetime and do not consider the effects of decreasing SoH on C_{eff} , further increasing the errors.

TABLE II. BATTERY MODEL VALIDATION

Battery	Error
Li-Ion ₅	4.35% ± 2.05%
Li-Ion ₆	5.83% ± 3.60%
Li-Ion ₇	3.84% ± 2.75%

We validate our model using the battery data available from the NASA Ames Prognostics Data Repository [26]. The repository includes the measurements of 2Ah Li-ion batteries charging and discharging at different currents and temperatures. Each measurement has the complete charge/discharge profile of a single battery until the end-of-life condition. We check our model using the results of three batteries tested at room temperature. Table II shows that our model has 4.67% average error compared to the measurements.

In summary, a battery-based peak shaving architecture needs to have 1) an appropriate power delivery option and battery placement design to eliminate the power losses as much as possible, 2) finer grained control of the battery power output to shave longer peaks, 3) low battery coordination overhead to ensure timely response to peak pulses, 4) a cost effective battery type to reduce the battery costs and a battery model to accurately estimate the effects of high discharging current.

IV. OUR PROPOSED ARCHITECTURE

In previous sections, we show that previous designs do not capture the effects of a binary battery discharge, which requires high discharging current. Furthermore, the distributed design requires a centralized controller to obtain its peak performance. The performance of this controller highly depends on the data center interconnect and can be negatively affected with increasing delay. Finer-grain control with the smaller batteries is a key to achieving good peak shaving results. This section presents our battery-based peak shaving architecture. Our design places the batteries in a centralized location and connects their aggregate output to the utility grid with a grid-tie inverter. Our model outperforms the distributed design by exploiting the

nonlinear nature of Peukert's Law, despite additional DC-AC conversion losses in the centralized UPS. It obtains improved battery lifetime and requires significantly less communication overhead than SoA distributed designs.

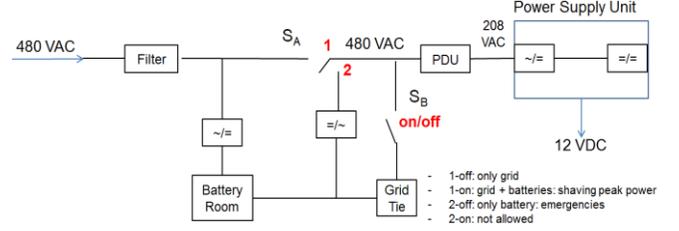


Figure 6: Our battery placement proposal using a grid-tie

Instead of decentralizing the batteries, we place them together and connect the batteries to the main grid using a grid-tie inverter. A grid-tie inverter allows any quantity of DC power to be converted into AC and fed into the grid in an efficient way [17], [27]. Figure 6 presents the layout of our design. The integration of battery power to the system is controlled with two switches. When switch S_A is in mode "1", the data center operates normally, accepting any quantity of battery power output from the grid-tie, which is controlled by the switch S_B . If it is "OFF", the grid is the only power supplier, i.e. we are under the power threshold. If it is "ON", the batteries are active and shaving peak power. The batteries can be recharged directly by the grid through a rectifier. S_A is in mode "2" only in emergency cases, making sure that the only power supplier is the battery. The case where S_A is in mode "2" and S_B is "ON" is not allowed because it just combines the same battery output.

Even though the batteries are centralized, we still treat them as distributed and enable them to individually charge/discharge. The fact that grid-tie inverter allows any quantity of DC to be combined with AC makes it possible to adaptively select the discharge current of the batteries. Instead of using batteries with high current rates as in both SoA centralized and distributed designs, we can increase the number of batteries being discharged and scale down the current. In fact, this leads to finer grained control of the battery output compared to both existing designs. Furthermore, having more batteries used simultaneously with the same discharging current, we reduce the variation in battery discharge profiles. Decreasing both this variation and discharging current helps increase the battery lifetimes.

We place the batteries together and allow the discharging current to scale down instead of being in a binary mode. We have a set, Φ , of discharge currents, and we choose the smallest current from Φ that can sustain the peak demand with the available batteries. Based on this current, we compute the number of batteries to use:

$$I_d = \min_{I \in \Phi} \{ I \mid V_b * I * N_b > (P_d - P_t) \ \& \ N_b \leq N_a \} \quad (5)$$

$$N_b = \left\lceil \frac{(P_d - P_t)}{V_b * I} \right\rceil \quad (6)$$

where V_b is the voltage of a single battery, P_d is the peak power demand, P_t is the peak power threshold to sustain, N_a is the number of available batteries, N_b is the number of

batteries required to discharge and I_d is the selected discharging current. These equations make sure that the minimum feasible discharging current is selected over all the selected batteries by ensuring the number of selected batteries is smaller than the number of available batteries. The set of available batteries include all the batteries having SoC greater than $100 - DoD_{goal}$, where DoD_{goal} is a predetermined value between 1 and 100 to better control the battery lifetime [5], [4]. Larger DoD_{goal} values can shave bigger peak power pulses for a longer duration but they lead to shorter average battery lifetime values. We refer this process as the *discrete current* policy.

This policy may select a subset of batteries to discharge. During battery selection, we choose the batteries available with the greatest SoH values. This minimizes the probability that a battery breaks down during discharging and it is the best a controller can do without any knowledge about the future power demand. The advantage of our architecture is that since the batteries are placed centrally they do not need to go through the data center network to coordinate for the battery selection process. They can use a dedicated network for this coordination. Thus, we can use a centralized controller with a much smaller expected latency, up to 4x less vs. the distributed design.

Alternatively, we can use all the available batteries to discharge at the same current. We define the number of the available batteries, i.e. the ones with SoC greater than $100 - DoD_{goal}$, as N_a . The discharging current, I_d , becomes:

$$I_d = \left\lceil \frac{(P_d - P_{th})}{V_b * N_a} \right\rceil \quad (7)$$

where P_d is the power demand, P_t is the peak power threshold and V_b is the voltage of a single battery. Different than the previous policy, this policy does not let any battery to be idle during a peak power pulse, i.e. a battery is either drained or discharging. As a result, it does not have a predefined set of discharging currents and it selects the discharging current on-the-fly based on the number of available batteries. We refer this process as the *all battery* policy. Since it discharges all the available batteries, there is no battery selection problem.

We use AC power delivery because it is most common in today's data centers and existing systems can apply our design without new infrastructure cost. Despite the power losses associated with the centralized placement, we still use it because of its simplicity and low maintenance requirements. We address this problem by adding 8% (section VII.C) more batteries into our architecture and compensating the additional capacity cost with elevated battery life. Furthermore, our design can leverage a dedicated network to establish coordination among the batteries, instead of being dependent on the data center network, reducing the communication overhead.

We compare our design against SoA designs in Table III in terms of the key architectural challenges we describe previously. Our design leverages the useful properties of existing designs that are necessary to shave long peaks. We add the ability to adjust the discharging current adaptively and a detailed battery model to capture the effects of a high discharge current. Also, our design can facilitate the locality

of the batteries by using a dedicated network to establish the communication, instead of using the data center network.

TABLE III. COMPARISON OF OUR DESIGN VS. STATE-OF-THE-ART (SOA)

	SoA [2] centralized	SoA [4], [5] distributed	Our design
Placement	Centralized	Distributed	Centralized
Selective Discharge	X	✓	✓
Adaptive Current	X	X	✓
Battery Model	X	X	✓
Coordination Medium	N/A	Data center network	Dedicated network

V. COST MODELS

This section presents the cost models used by the previous work to quantify the benefits of the peak power shaving. For each different model, we show the domains they are applicable to, how they are calculated and specifically focus on how the battery cost affects the overall cost. The latter part is important since we show that the average battery lifetime is overestimated by up to 2.44x with simplistic battery models, increasing the battery cost with more frequent replacements. As a result, the benefits of peak shaving with batteries are also overestimated.

A. Co-location Rental (CLR) Cost Model

Co-location providers rent their data center equipment and space to retail customers. This applies to companies that require a data center-like system but do not want to build their own. A well-known example for a co-location renter is content delivery networks (CDNs) [4]. These renters make long-term power contracts with co-location providers and pay based on their provisioned power, instead of their actual consumed power. As a result, decreasing their peak power consumption immediately translates to savings (Figure 1-b). Palasamudram et al. [4] target this domain for their distributed battery-based peak shaving design and calculate the total cost as:

$$Cost_{total} = c_p * PP_{total} + \frac{c_b}{L} * BC_{total} \quad (8)$$

where c_p is the unit power price, PP_{total} is the total provisioned power, c_b is the unit battery price, BC_{total} is the total battery capacity and L is the expected battery lifetime. Then, they calculate the savings as:

$$Savings = 100 * \frac{Cost_{total}(no\ batteries) - Cost_{total}(batteries)}{Cost_{total}(no\ batteries)} \quad (9)$$

where $Cost_{total}(batteries)$ and $Cost_{total}(no\ batteries)$ represents total cost with and without batteries, respectively. When calculating the total cost without the batteries, we can just neglect the battery related parts of Equation 8. The main purpose of peak shaving in this case is to reduce the provisioned power level so that the co-location renters can contract for less power.

B. Total Cost of Ownership (TCO) Model

There are several companies that own their data centers, where they still make power contracts based on their peak power consumption to reduce their cost of energy. However,

they achieve this peak value rarely and underutilize the provisioned power. A solution to this is to add more servers to the data center, which improves the power utilization but also increases the peak power level. A peak shaving mechanism can ensure that the provisioned power level is not violated with additional servers. In this case, the provisioned power level does not decrease but both the provisioned power and the data center equipment can be used to host more servers and thus TCO/server reduces. Also, assuming that each server brings a constant amount of revenue, the total profit increases [5]. This also shows that the savings is directly proportional to TCO/server reduction.

TABLE IV. TCO MODEL INPUTS RELATED TO THE BATTERIES

Input	LA Value	LFP value
Battery unit price -rated with 20h	2 \$/Ah [30]	5 \$/Ah [29]
Per server capacity	20 Ah [5]	40 Ah [5]
Peukert's exponent	1.15 [23]	1.05 [23]
Battery nominal voltage	12V [11]	
Data center depreciation time	10 years [31]	
Server depreciation time	4 years [31]	
Utility energy price	4.7¢/kWh [32]	
Utility power price	12 \$/kW [2]	

Kontorinis et al. [5] use this analysis by collecting the depreciation and opex data from the APC's commercial TCO calculator [28]. This model computes the TCO/server by dividing it into multiple parts, calculating each part separately and analyzing how each part changes with more servers within the same power budget. Table V summarizes the different components of TCO and shows the TCO break-down for different designs. More servers decrease the TCO/server and increase the profit obtained from a server. We compare the TCO/server of each battery placement design in our study with the TCO/server of a data center which does not use batteries for peak shaving (base model). The part we are interested in TCO partitioning is the *UPS depreciation*, accounting for the battery costs. If the associated UPS depreciation cost is high, we can obtain negative savings compared to the base model. Some reasons for high UPS depreciation include short average battery lifetime (requires frequent replacements) or using an inappropriate battery type for peak shaving (low energy density, short service time, etc.). Table IV lists the input

values for both this model and CLR model. Further details of the TCO model are not in the scope of this paper and covered in detail in [5].

Our grid-tie design requires more power distribution infrastructure than the distributed design because we keep transmitting power throughout the data center, even if the power is not drawn from the utility. For example, a 10MW data center may have 1MW worth of additional servers due to peak shaving. In our case, the extra power is provided from the UPS through the data center power infrastructure to the servers. In the distributed case, this extra power is not provided through the data center power infrastructure. Although all the servers are connected to the main power infrastructure, during a peak pulse some of them may disconnect themselves from the main power infrastructure and get power locally from the on-board UPS. Therefore, the provisioned power infrastructure is sufficient. This means that our approach has constant power infrastructure depreciation, whereas the distributed design decreases this depreciation with more servers. But, our design does not require a custom PSU or power distribution, as opposed to the DC architecture. This makes it practical for the existing data centers. The additional peak shaving opportunities from our approach outweigh the additional infrastructure costs.

VI. METHODOLOGY

Power Measurements and Workloads: We estimate the power consumption of a large scale data center using measurements from our data center container. It has 200 servers consisting of Nehalem, Xeon and Sun Fire servers running Xen VM. We compose a mix of common benchmarks to measure power and performance of different jobs on our servers. We use RUBiS [33] to model service-sensitive, eBay-like, workloads with 90th%ile of response times at 150ms, and Olio [34] to model social networking workloads with response times ranging from 100ms up to multiple seconds depending on the type of data uploaded (e.g. text vs. photos). Multiple Hadoop [35] instances are run as batch jobs with 2 min mean arrival time and with average execution time of 10 min [36]. Performance is measured at 10ms sampling rate, while power is obtained at 60Hz. We create an event-driven simulator embedding the power

TABLE V. TCO/SERVER BREAK-DOWN IN DIFFERENT DESIGNS [5]. THE COMPONENTS WITH DIFFERENT TRENDS ARE HIGHLIGHTED.

TCO Component	w/o peak shaving	Distributed Design Break-down			Grid-tie Design	
		TCO/server trend with more servers	Battery Model		TCO/server trend with extra servers	Break-down
			Simple	Detailed		
Facility space depreciation	\$3.40	Decreasing	\$2.74	\$2.74	Decreasing	\$2.72
UPS depreciation	\$0.13	Constant	\$1.67	\$5.00	Constant	\$3.33
Power infrastructure depreciation	\$5.94	Decreasing	\$4.79	\$4.79	Constant	\$5.94
Cooling infrastructure depreciation	\$2.46	Decreasing	\$1.98	\$1.98	Decreasing	\$1.96
Racks, monitoring, installation	\$8.97	Decreasing	\$7.23	\$7.23	Decreasing	\$7.17
Data center opex	\$7.49	Decreasing	\$6.04	\$6.04	Decreasing	\$5.99
Server depreciation	\$31.25	Constant	\$31.25	\$31.25	Constant	\$31.25
Server opex	\$1.56	Constant	\$1.56	\$1.56	Constant	\$1.56
PUE overhead	\$1.94	Constant	\$1.94	\$1.94	Constant	\$1.94
Utility monthly energy cost	\$8.71	Constant	\$8.71	\$8.71	Constant	\$8.71
Utility monthly power cost	\$4.20	Decreasing	\$3.39	\$3.39	Decreasing	\$3.36
Total	\$76.04	Decreasing	\$71.30	\$74.63	Decreasing	\$73.94

information and the workload characteristics of the measurements to simulate a data center. We model each 8-core server with an M/M/8 queue, and use a linear CPU utilization based power estimate commonly used by others [1], [6]. Table VI shows that the average simulation error is below 10% for all quantities of interest, with 3% for power estimates, while performance for services has only 6% and MapReduce completion times are within 8% of measured values on our data center container.

TABLE VI. VERIFICATION OF POWER AND PERFORMANCE MODELS

Parameter	Average Error
Avg. Power Consumption	3%
90 th tile Services QoS	6%
Avg. MapReduce Comp. Time	8%

TABLE VII. WORKLOAD PARAMETERS

Workload	Average Time	
	Service	Interarrival
Search [6]	50ms	42ms
Social Networking [34]	1sec	445ms
MapReduce [37]	2 min	3.3 min

We model typical user request load onto a full data center to understand the benefits of peak power shaving. We use a year of publicly available traffic data of two Google products, Orkut and Search, as reported in Google Transparency Report [38] to represent latency-centric service jobs. We reproduce the weekly waveform of MapReduce jobs shown in Figure 3 of [37] to model batch jobs. Table VII shows the parameters of these workloads. We created a week’s worth of workload mixture. Figure 7 compares the workload components to the maximum load. The max load ratio is around 80% while the average is 45%, shown by the horizontal line.

Data Center and Battery Simulation: We limit our data center simulation to a week because it is not computationally feasible over long periods due to fine event granularity. We extract the power profile of the data center as well as the charge/discharge profile of the batteries in the given timeframe and scale these profiles appropriately for longer time intervals. We refer to this process as data center workload simulation. The main goal of this preprocessing is to analyze the required DoD level and discharging current profiles for the batteries.

Figure 8 shows the DoD level variation of the grid-tie design and the distributed design with different level controllers over a week using LFP batteries when DoD_{goal} is set to 60%. Both designs shave 15% of the peak power. The grid-tie architecture is more consistent, followed by high level distributed controllers. In these cases, the batteries use all the available capacity, because the battery power is distributed evenly across the batteries. In contrast, the DoD value is distributed between 20% and 60% approximately uniformly with a server level controller since individual server power profiles vary and there is no coordination between them. In Figure 9, we present the average discharging current profile of the distributed and grid-tie design over a 3 day period from the same experiment described above. The grid-tie design reduces the discharging

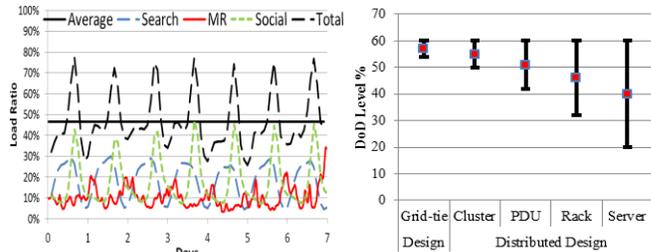


Figure 7. Data center workload mix

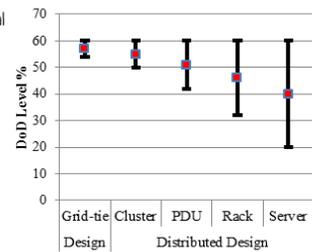


Figure 8. DoD level variation

current significantly without affecting the amount of peak power shaved, and thus can decrease the negative effects of high discharging current.

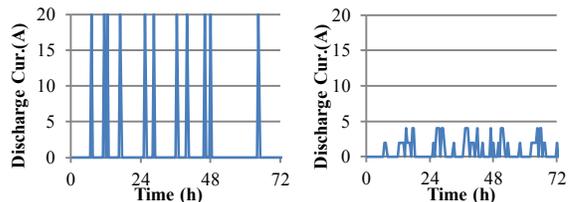


Figure 9. Avg. discharging current for the distributed design (left) and grid-tie design (right) over a 3 day period, with LFP.

We include both LFP and LA batteries in our study and assume that the battery capacity per server is 40Ah and 20Ah respectively with 12V nominal voltage. These capacity values are the maximum that can fit into a 2U server [5]. Each battery is allowed to discharge up to DoD_{goal} . We change the DoD_{goal} to see how it impacts both average battery lifetime and peak power level that can be sustained. Our battery model estimates the SoC and SoH of each battery. After analyzing short-term battery usage profiles, we use our battery model and simulate only charge/discharge cycles of the batteries. We run the simulation for several years of simulation time to estimate the battery lifetime. We consider a battery dead when its SoH goes below 80% [24], [25]. We refer to this process as battery simulation.

VII. RESULTS AND EVALUATION

This section first shows the effects of incorporating a detailed battery model on savings. We analyze CLR and TCO cost models and show that the savings are overestimated with both by up to 3.35x. Then, we compare the peak shaving capabilities and efficiencies of different designs. DC distribution gives the best performance, but it is not preferable for current data centers as it requires a complete redesign of the power delivery equipment. Finally, we compare our design with SoA designs in terms of average battery lifetime and cost savings. Our design achieves up to 78% longer battery lifetime and thus, up to 50% more cost savings with only 8% more batteries to account for non-ideal battery characteristics.

TABLE VIII. BATTERY LIFETIME ESTIMATION COMPARISON

	LA	LFP
SoA low current rated estimates [4], [5]	3 years	10 years
Our estimates	1.4 years	4.1 years

TABLE IX. CLR COST SAVINGS FOR DISTRIBUTED LA AND LFP BATTERIES

LA – distributed design					LFP – distributed design				
Cost	2 \$/Ah	3 \$/Ah	4 \$/Ah	5 \$/Ah	Cost	5 \$/Ah	8 \$/Ah	10 \$/Ah	12 \$/Ah
Lifetime	CLR Savings (%)				Lifetime	CLR Savings (%)			
1	0.9	-3.6	-8.1	-12.7	1	-25.3	<-50	<-50	<-50
2	5.5	3.2	0.9	-1.3	2	-2.7	-16.3	-25.3	-34
3	6.4	4.6	2.7	1	3	1.8	-9.4	-17	-24
4	8.2	5.9	4.6	3.2	4	6.4	-1.8	-7.2	-13
7	Not possible				7	11	5.5	1.9	-1.8
10					10	15.5	12.7	11	9.1

TABLE X. TCO/SERVER SAVINGS FOR DISTRIBUTED LA AND LFP BATTERIES

LA – distributed design					LFP – distributed design				
Cost	2 \$/Ah	3 \$/Ah	4 \$/Ah	5 \$/Ah	Cost	5 \$/Ah	8 \$/Ah	10 \$/Ah	12 \$/Ah
Lifetime	TCO/server Savings (%)				Lifetime	TCO/server Savings (%)			
1	0.02	-2.08	-4.18	-6.27	1	-13.48	-26.63	-35.4	-44.17
2	2.21	1.2	0.2	-0.81	2	-2.52	-9.1	-13.48	-17.87
3	2.65	1.86	1.07	0.28	3	-0.33	-5.59	-9.1	-12.61
4	3.09	2.52	1.95	1.38	4	1.86	-2.08	-4.71	-7.35
7	Not possible				7	4.05	1.42	-0.33	-2.08
10					10	6.24	4.93	4.05	3.18

A. Effects of the Detailed Battery Model

We start our evaluation by quantifying the effects of a simplistic battery model on the cost. We use our battery model with our long term battery simulator to estimate the average lifetime of an LA and LFP battery when shaving peak power. Table VIII shows the comparison between our long-term battery lifetime estimates and the low-current rated estimates used by the previous work. We see that neglecting the effects of high current results in 2.14x and 2.44x longer battery lifetime predictions. This leads to underestimated battery costs and overestimated cost savings. Table IX and X show the CLR and TCO/server savings, for both LA and LFP batteries with varying battery lifetime values. In our analysis, we obtain 9.5% and 19% peak power capping with distributed LA and LFP batteries, respectively. This peak shaving also enables 10.5% and 24% extra servers to be deployed within the same power budget when using the TCO model. We first use inexpensive batteries rated at low currents. In this case, CLR savings are 2.8% and 6.4%; TCO/server savings are 0.9% and 1.86% for LA and LFP batteries, respectively. If we do not capture the effects of high discharge current, these savings are 6.4% and 15.15% for CLR model and 2.65% and 6.24% for the TCO model. The savings are overestimated by up to 2.94x and 3.35x for LA and LFP.

Next, we use batteries with larger rated currents: 10h, 5h and 1h [39], that are also more expensive: 8, 10, 12\$/Ah for LFP and 3, 4, 5 \$/Ah for LA. The average LFP lifetime increases to 5, 6 and 8 years and 2, 2.5 and 3 years for LA. Table IX shows that CLR cost savings become 3.2%, 2.7%, 1% for LA and 5.5%, 1.9%, -1.8% for LFP batteries with 10h, 5h and 1h rated batteries. Similarly, Table X shows that TCO/server savings become 1.2%, 0.94%, 0.28% for LA and 1.42%, -0.33% and -2.08% for LFP with 10h, 5h and 1h rated batteries respectively. Although the battery lifetime values are closer to the low current rated estimates, higher battery price overshadows the savings obtained by fewer battery replacements.

B. Peak Shaving Efficiency of State-of-the-Art Designs

We continue our evaluation by comparing the peak shaving capabilities of SoA battery placement designs. We also include our battery model to account for the high discharge currents. Most data centers use a centralized LA battery, powering the entire data center when it is active and not overprovisioned for peak shaving. The capacity of this battery is adjusted to handle only emergency cases, which last a few minutes. We assume that this design has 3200 Ah worth of LA batteries [2], [5]. Then, we compute the amount of time a battery can shave a fixed peak pulse and how long it takes to fully recharge it during low demand. Table XI shows that the centralized design shaves the peak power for only 3-4 minutes when $P_{threshold}$ is set to 255W per server. It cannot sustain long peaks and needs to apply other policies such as DVFS which have performance overhead.

TABLE XI. PEAK SHAVING CAPABILITIES OF THE CENTRALIZED DESIGN WITH DIFFERENT BATTERY TYPES. $P_{THRESHOLD}$ IS SET TO 255W PER SERVER.

Peak Power Per Server (W) – Shaving %	Power Shaving Duration (min)		
	Centralized – LA not scaled	Centralized – LA scaled	Centralized – LFP scaled
300 – 15%	3.81	24.24	70.46
310 – 17.5%	3.67	23.34	68.07
320 – 20.3%	3.54	22.50	65.84

To address this problem, we increase the capacity of the centralized battery by 5x and obtain ~6x longer peak shaving. LA batteries have large volume, so the capacity cannot be scaled significantly. The increase in peak shaving duration is more than 5x because the stress on discharging current rate decreases nonlinearly as a result of Peukert’s law [13]. In contrast, the recharging duration increases almost linearly with scaling capacity. The peak shaving duration, despite increased capacity, is still not sufficient enough to sustain peaks lasting hours. Another option is to use LFP batteries with more energy density and less nonlinear battery behavior [5]. This can scale up the capacity further. We use a

TABLE XIII. EFFICIENCY OF CENTRALIZED VS. DISTRIBUTED DESIGNS CONSIDERING DIFFERENT POWER EQUIPMENT AND DELIVERY OPTIONS

Unit	Efficiency common	Efficiency best	Design	% Battery energy wasted before providing server power		% Grid energy wasted before charging the batteries	
				Common	Best	Common	Best
Centralized double conversion UPS	85% [47]	90% [19]	Distributed w/ AC power	5%	2%	35%	23%
AC distribution PDU	98% [19]						
Server AC PSU	75% [46]	90% [19]	Distributed w/ DC power	2%	2%	12%	8%
DC UPS	92% [21], [47]	95% [19]					
DC distribution PDU	99% [19]		SoA Centralized	38%	15%	5%	5%
Server DC PSU	92% [19]						
Filter + Rectifier	95% [19]	97% [48]					
Grid-tie inverter	95% [45]						

total capacity of 40K Ah [5] and get up to 70 minutes of peak shaving at high cost. The peak shaving benefits are insufficient to compensate for high battery costs. This analysis shows that centralized battery design is not a good option for peak shaving when the battery powers the entire data center in boolean fashion as in the state-of-the-art work.

TABLE XII. PEAK SHAVING AND BATTERY RECHARGING CAPABILITIES OF THE DISTRIBUTED DESIGN WITH DIFFERENT BATTERY TYPES AND AC VS. DC POWER DELIVERY OPTIONS. $P_{THRESHOLD}$ IS SET TO 255W PER SERVER.

Peak Power Per Server (W) – Shaving %	Power Shaving Duration (min)		
	Distributed – LA with AC	Distributed – LFP with AC	Distributed – LFP with DC
300 – 15%	192.92	552.23	552.23
310 – 17.5%	157.07	451.09	451.09
320 – 20.3%	132.27	381.08	381.08
Low Power Per Server (W)	Recharging Duration (h)		
	Distributed – LA with AC	Distributed – LFP with AC	Distributed – LFP with DC
220	8.47	16.95	14.81
210	6.58	13.16	11.56
200	5.38	10.75	9.43

Ideally, batteries should supply power only for the portion above the peak level. The centralized design cannot achieve this because it operates in boolean mode at data center level. The distributed design allows battery power to be controlled in finer granularity by selectively discharging only a subset of all the batteries. We analyze the power shaving duration of distributed LFP and LA batteries in Table XII. The size of each LA and LFP battery is set to be 20Ah and 40Ah, respectively. These are the maximum capacities that can fit in a 2U server [5]. Although the LFP capacity is more than LA by 2x, it shaves a given peak for ~3x longer because LFP battery behavior is less nonlinear at high current, proving to be a better fit for the distributed design. But, recharging all the batteries back to back takes more time for LFP due to its larger capacity. Since batteries can selectively discharge, this is not much of an issue.

Another important key challenge is to reduce the conversion losses that impact the effective battery input/output. The distributed design puts the batteries next to the servers and increases the effective battery capacity compared to the centralized design. DC power delivery can be used to further eliminate the conversion losses on the power path, reducing the input power required to recharge the battery. We show the best and common efficiency values for the power infrastructure of both AC and DC options in

Table XIII. It also shows the amount of energy wasted to recharge the batteries and battery output wasted before going into the servers.

The centralized design does not waste a lot of grid power but the battery output loss is 15%, which further reduces its peak shaving duration. We see that the distributed DC design obtains the best efficiency by having the smallest total conversion losses. The AC counterpart provides similar battery output power but it wastes the grid input 3x more than the DC design and results in longer discharges. Table XII also shows the comparison between AC and DC distributed options in terms of effective discharge and recharge durations. Discharging capabilities are the same but the DC design takes 14% shorter time to fully recharge, which makes it a safer option as batteries get ready for the next peak earlier. Although the DC option is more energy efficient, it is an unfeasible option for existing data centers because its high cost to replace the power infrastructure.

C. Performance of Our Grid-tie Design

We compare our grid-tie design with previous designs in terms of energy efficiency, average battery lifetime, cost savings, and communication overhead. As we place the batteries in a centralized location, we still lose 15% of battery output because of the conversion losses (see Table XIII). However, batteries are used at lower discharge current and have higher effective battery capacity. This reduces the effects of the conversion losses. Instead of 15% performance difference, we get an average of 8% performance loss compared to the distributed design as shown in Table XIV. We compensate for this performance loss by adding 8% more battery capacity, which is feasible because we are not limited by rack size as in the distributed design.

Table XV shows the power shaving statistics of our grid-tie design and the distributed design. We analyze our design with and without additional battery capacity as well as with *all_battery* and *discrete_current* policies (see section IV). The average battery lifetime does not change with additional battery capacity, but the *all_battery* policy results in longer average battery lifetime. The average battery lifetime estimates are 5.4 and 2.2 years for LFP and LA respectively using the *discrete_current* policy. We obtain 6.4 and 2.5 years with the *all_battery* policy. The battery lifetime values are 60% and 78% higher compared to the distributed design for LFP and LA batteries respectively since the discharging current can be scaled down with our design so that the negative impact on the battery lifetime is minimized. The

all_battery policy scales down the discharging current more by using all available batteries and thus performs better than the *discrete_current* policy.

TABLE XIV. PEAK SHAVING CAPABILITIES OF OUR DESIGN COMPARED TO THE DISTRIBUTED DESIGN. $P_{\text{THRESHOLD}}$ IS SET TO 255W PER SERVER.

Peak Power Per Server (W) – Shaving %	Power Shaving Duration (min)	
	Distributed – LFP	Grid-tie –LFP
300 – 15%	552	516
310 – 17.5%	451	418
320 – 20.3%	381	351

Our grid-tie design with 8% larger capacity obtains similar peak shaving performance compared to the distributed design. It compensates the increased battery costs with longer battery lifetime. Our design achieves up to 11% and 5.5% savings for LFP and LA batteries when renting from co-location providers. These savings are 70% and 100% higher than the distributed design. Similarly, we obtain up to 2.77% and 1.87% TCO/server savings using LFP and LA respectively. These TCO/server savings correspond to up to \$75K/month for a 10MW data center [28]. The TCO savings are 48% and 107% higher than the savings of the distributed design.

The distributed design requires a centralized controller to get the best peak shaving performance [4], [5]. Since the batteries are distributed to the servers, this controller communicates with the batteries through the data center interconnect. High network usage leads to large signal delays to/from batteries. This can affect the performance of the controller negatively by increasing the response time to a peak pulse or transmitting outdated battery and server load information. The distributed design can also use multiple controllers placed at different levels of power hierarchy [5]. A decentralized control mechanism significantly reduces the peak shaving capabilities (see Table I). Our design can isolate itself from the data center interconnect, achieving fast communication even with high network congestion.

Figure 10 compares the total delay of our grid-tie design during a discharge process with that of different controllers deployed in distributed design. We analyze the worst-case scenario where the controller needs to poll each battery. The left vertical axis is on a log scale and shows the communication delay whereas the right vertical axis presents the peak shaving percentage achieved by each configuration. We assume a fat-tree network topology [40] and model the links in the network with 10 Gbps capacity, transmitting a 1K package at 1us [41]. We evaluate three different network congestion levels: an ideal network, network with normal and high level congestion. The first one has no queuing delay

whereas the other two have 50 us and 350us delay when transmitting a single message in a switch, respectively [42]. In this experiment, cluster level corresponds to centralized communication for the distributed design. The low-level controllers have less total delay compared to our design in the ideal network case, but as the network congestion increases, our design performs better, except for the rack level controller, which has 60% less peak shaving performance than our design. Our design has similar peak shaving performance (1% better) compared to the centralized control in distributed design. But, even in the ideal case of network, our design has around 20 ms total delay compared to 100 ms of the centralized control for the distributed design. Even in this case, we obtain 4x less communication overhead, and this difference increases exponentially as the network delay ramps up.

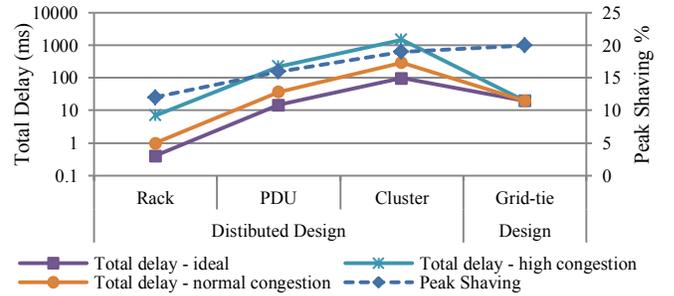


Figure 10. Communication and peak shaving performance of the grid-tie design vs. the distributed design

VIII. CONCLUSION

Peak shaving with batteries has gained significant importance because of its ease of applicability and great performance. In this paper, we address the key challenges of architecting a cost and energy efficient battery-based peak shaving design. We first use a detailed battery model to capture the negative effects of high discharging currents. Our results indicates that not having a detailed battery model overestimates the battery lifetime up to 2.44x and leads to 3.35x error in cost saving estimates. Second, we propose a new grid-tie based design which preserves the advantages of the existing designs, such as individual control of the batteries, and eliminates the key drawbacks, such as adaptively selecting the discharge current. It can use a fast, dedicated network to coordinate the batteries, reducing the communication overhead by 4x compared to the distributed design. Our design achieves up to 78% longer battery lifetime and doubles the savings compared to the state-of-the-art designs.

TABLE XV. GRID-TIE VS. DISTRIBUTED DESIGN, HIGHLIGHTED BEST VALUES. (EB = EXTRA BATTERIES, BL = BATTERY LIFETIME, PS = PEAK SHAVING, ES = EXTRA SERVERS)

Design - Policy	EB	LFP					LA				
		BL	PS	ES	CLR savings	TCO/server savings	BL	PS	ES	CLR savings	TCO/server savings
Grid-tie <i>all_battery</i>	8%	6.4	20%	25%	11%	2.77%	2.5 yrs	9.9%	11%	5.5%	1.87%
	0%	yrs	16%	19%	7.7%	1.36%		7%	8%	2.8%	1.14%
Grid-tie <i>discrete_current</i>	8%	5.4	20%	25%	9.4%	2.77%	2.2 yrs	9.9%	11%	4.9%	1.44%
	0%	yrs	16%	19%	6.1%	1.36%		7%	8%	2.2%	0.42%
Distributed -N/A	N/A	4 yrs	19%	24%	6.4%	1.86%	1.4 yrs	9.5%	10.5%	2.7%	0.9%

ACKNOWLEDGEMENTS

This work was sponsored in part by Google, NSF ERC CIAN (grant number 812072), NSF IRNC TransLight–StarLight (grant number 962997), and CNS. The authors also acknowledge the support of the Multiscale Systems Center (MuSyC), one of six centers under the Focus Center Research Program (FCRP), a Semiconductor Research Corporation program (SRC).

REFERENCES

- [1] X. Fan, W. Weber and L. Barosso, "Power provisioning for a warehouse-sized computer," in *ISCA* 2007.
- [2] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, "Benefits and Limitations of Tapping into Stored Energy For Datacenters," *ISCA* 2011.
- [3] S. Govindan, D. Wang, A. Sivasubramaniam and B. Urgaonkar, "Leveraging stored energy for handling power emergencies in aggressively provisioned datacenters," in *ASPLOS*, 2012.
- [4] D. Palasamudram, R. Sitarman, B. Urgaonkar and R. Urgaonkar, "Using Batteries to Reduce the Power Costs of Internet-scale Distributed Networks," in *Symposium on Cloud Computing*, 2012.
- [5] V. Kontorinis, L. Zhang, B. Aksanli, J. Sampson, H. Houman, E. Pettis, D. Tullsen and T. Rosing, "Managing Distributed UPS Energy for Effective Power Capping in Data Centers," in *ISCA* 2012.
- [6] D. Meisner, C. Sadler, L. Barroso, W. Weber and T. Wenisch, "Power management of online data-intensive services," in *ISCA* 2011.
- [7] R. Nathuji and K. Schwan, "Vpm tokens: virtual machine-aware power budgeting in datacenters," in *HPDC* 2008.
- [8] N. Buchbinder, N. Jain and I. Menache, "Online job-migration for reducing the electricity bill in the cloud," in *in Proc. Networking* 2011.
- [9] 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 *INFOCOM*, 2010.
- [10] Y. Zhang, Y. Wang and X. Wang, "Capping the electricity cost of cloud-scale data centers with impacts on power markets," *HPDC*, 2011.
- [11] Google, "Google Summit," 2009.
<http://www.google.com/corporate/datacenter/events/dc-summit-2009.html>.
- [12] Facebook, "Hacking conventional computing infrastructure," 2011.
<http://opencompute.org/>.
- [13] SmartGauge, "Peukert's law equation and its explanation," 2011.
<http://www.smartgauge.co.uk/peukert.html>.
- [14] D. Meisner, B. Gold and T. Wenisch, "PowerNap: Eliminating Server Idle Power," in *ASPLOS* 2009.
- [15] D. Wang, C. Ren, A. Sivasubramaniam, B. Urgaonkar and H. Fathy, "Energy Storage in Datacenters: What, Where, and How much?," in *SIGMETRICS* 2012.
- [16] S. Drouilhet and B. Johnson, "A Battery Life Prediction Method for Hybrid Power Applications," in *AIAA Aerospace Sciences Meeting and Exhibit*.
- [17] N. Deng, C. Stewart and J. Li, "Concentrating Renewable Energy in Grid-Tied Datacenters," in *ISSST* 2011.
- [18] A. Pratt, P. Kumar and T. Aldridge, "Evaluation of 400V DC distribution in telco and data centers to improve energy efficiency," in *INTELEC* 2007.
- [19] M. Ton, B. Fortenbery and W. Tschudi, "DC Power for Improved Data Center Efficiency," Lawrence Berkeley National Laboratory, 2008.
- [20] L. Benini, G. Castelli, A. Macii, E. Macii, M. Poncino and R. Scarsi, "Discrete-time battery models for system-level low-power design," *IEEE Trans. on VLSI Sys.* 2001.
- [21] D. Rakhmatov, S. Vrudhula and D. A. Wallach, "Battery lifetime prediction for energy-aware computing," in *ISLPED* 2002.
- [22] K. Soon Ng, C.-S. Moo, Y.-P. Chen and Y.-C. Hsieh, "Enhanced coulomb counting method for estimating state-of-charge and state-of-health of lithium-ion batteries," *Applied Energy*, vol. 86, no. 9, 2009.
- [23] F. Harvey, "Table with Peukert's exponent for different battery models," 2001.
http://www.electricmotorsport.com/store/ems_ev_parts_batteries.php.
- [24] PowerSonic, "Technical manual of LA batteries,"
<http://www.power-sonic.com/technical.php>.
- [25] P. Rong and M. Pedram, "An analytical model for predicting the remaining battery capacity of lithium-ion batteries," *IEEE Transactions on VLSI Systems* 2006.
- [26] B. Saha and K. Goebel, "Battery Data Set, NASA Ames Prognostics Data Repository," 2007.
<http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/>.
- [27] H. Qian, W. Yu, J.-S. Lai and J. Zhang, "A High-Efficiency Grid-Tie Battery Energy Storage System," in *IEEE TPEL* 2011.
- [28] APC, "InfraStruxure Total Cost of Ownership, Infrastructure cost report," 2008. <http://www.apc.com/tools/isx/tco/>.
- [29] A. P. P. Corp, "Portable Power Product design, assemble and quality control," 2000. <http://www.batteryspace.com/lifepo4cellspacks.aspx>.
- [30] E. motor sport, "EV construction, thundersky batteries," 2001.
http://www.electricmotorsport.com/store/ems_ev_parts_batteries.php.
- [31] L. A. Barosso and U. Holzle, "The data center as a computer: An Introduction to the Design of Warehouse-Scale Machines, 2009.
- [32] D. E. Carolinas, "Utility bill tariff," 2009.
<http://www.duke-energy.com/pdfs/scschedulesopt.pdf>.
- [33] "RUBiS," <http://rubis.ow2.org/>.
- [34] Apache, <http://incubator.apache.org/olio/>.
- [35] "Hadoop," <http://hadoop.apache.org/>.
- [36] B. Aksanli, J. Venkatesh, L. Zhang and T. Rosing, "Utilizing green energy prediction to schedule mixed batch and service jobs in data centers," in *HotPower* 2011.
- [37] Y. Chen, A. Ganapathi, R. Griffith and R. Katz, "The case for evaluating MapReduce performance using workload suites," In Tech. Report No. UCB/EECS-2011-21, 2011.
- [38] "Google Transparency Report,"
<http://www.google.com/transparencyreport/traffic>.
- [39] "Battery prices at different current rates,"
<http://www.batteryspace.com/>.
- [40] D. Kliazovich, P. Bouvry, Y. Audzevich and S. Khan, "GreenCloud: A Packet-Level Simulator of Energy-Aware Cloud Computing Data Centers," in *GLOBECOM 2010*
- [41] M. Alizadeh, A. Kabbani, T. Edsall, B. Prabhakar, A. Vahdat and M. Yasuda, "Less is More: Trading a little Bandwidth for Ultra-Low Latency in the Data Center," in *USENIX NSDI 2012*
- [42] "Priority flow control: Build reliable layer 2 infrastructure,"
http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9670/white_paper_c11-542809.pdf.
- [43] M. Swierczynski, R. Teodorescu and P. Rodriguez, "Lifetime investigations of a lithium iron phosphate (LFP) battery system connected to a wind turbine for forecast improvement and output power gradient reduction," in *In BatCon'08.*, 2008.
- [44] Windsun, "Lead-acid batteries: Lifetime vs Depth of discharge," 2009.
http://www.windsun.com/Batteries/Battery_FAQ.htm.
- [45] US Department of Energy, "Grid-tie Energy Efficiency," 2010.
www1.eere.energy.gov/solar/review_meeting/pdfs/prm2010_apollo.pdf.
- [46] C. S. Computing. www.climatesaverscomputing.org/resources/certification.
- [47] Energy Star, "Uninterruptible Power Supply Energy Efficiency Values," www.energystar.gov/index.cfm?c=specs.uninterruptible_power_supplies.
- [48] Emerson Network Power, "Rectifier Energy Efficiency,"
http://www.emersonnetworkpower.com/en-US/Brands/EnergySystems/Pages/ensys_eSureRectifiers.aspx.