Exploring Highly Dependable and Efficient Datacenter Power System Using Hybrid and Hierarchical Energy Buffers

Longjun Liu, *Member, IEEE*, Hongbin Sun, *Member, IEEE*, Chao Li, *Member, IEEE*, Tao Li, *Fellow, IEEE*, Jingmin Xin, *Senior Member, IEEE*, and Nanning Zheng, *Fellow, IEEE*

Abstract—The massive and irregular load surges challenge datacenter power infrastructures. As a result, power mismatching between supply and demand has emerged as a crucial availability issue in modern datacenters which are either under-provisioned or powered by intermittent power sources. Recent proposals have employed energy storage devices such as the uninterruptible power supply (UPS) to address this issue. However, current approaches lack the capacity of efficiently handling the irregular and unpredictable power mismatches. In this paper, we propose Hybrid and Hierarchical Energy Buffering (HHEB), a novel heterogeneous and adaptive scheme that could enable various energy storage devices (ESDs) to be efficiently integrated into existing datacenters for dynamically dealing with power mismatches. Our techniques exploit the diverse characteristics of different ESDs and intelligent load assignment algorithms to improve the dependability and efficiency of datacenter power systems. We evaluate the HHEB design with a prototype. Compared with a homogenous battery energy buffering system, HHEB could improve energy efficiency by 39.7 percent, extend UPS lifetime by 4.7X, promote energy availability by 3.2X, reduce system downtime by 41 percent, and effectively improve the energy availability of various energy buffers in different hierarchies. It allows datacenters to adapt to various power supply anomalies, thereby improving operational efficiency, dependability and availability.

Index Terms—Datacenters, power management, energy storage, efficiency, dependability

1 Introduction

Power-related costs, including capital (CAP-EX) and operating (OP-EX) expenses, have become a significant fraction of Total Cost of Ownership (TCO) of datacenters. It is predicted that the power consumption of world datacenters alone will approach 1,000TWh within a decade (2013-2025), which is more than the total power now used for all purposes by Japan and Germany combined [1]. The huge power demands not only imply significant electricity cost expenditures but also lead to tremendous carbon emission. Therefore, Industry and academia alike are focusing more on the new perspectives of improving datacenter power efficiency and costs. Currently, there are two primary techniques: (1) aggressively under-provisioned datacenters power infrastructures (a.k.a., power under-provisioned datacenters), which has been proved as a meaningful methodology to dramatically reduce infrastructure capital expenditure (CAP-EX) and monthly recurring operating expenditure (OP-EX) [2], [3], [4], [5], [6], [7], [8]; (2) renewable energy integration into datacenter facilities. To effectively reduce carbon emission, not only academia has started to study the intermittent renewable energy power management schemes [9], [10], [11], [12], [13], [14], [15], [16], [17], [18], but also many IT companies, such as Microsoft, Apple, Google, etc., have begun to build renewable energy powered datacenters [19], [20], [21], [22], [23].

The above two power provisioning schemes can significantly reduce electricity cost and carbon emission. However, the power dependability and efficiency issues caused by power mismatches are more deteriorative, since (1) power under-provisioned datacenters intentionally subscribe lower power supply infrastructures, which may lead to power budget violations due to the irregular and bursty service requests, and (2) the nature of renewable power sources is intermittent and fluctuated, and it may exceed (i.e., valley power) or lower (i.e., peak power) than power demands even if the latter are stable.

We classify existing proposals of handling the power mismatching issue into two categories: (1) performance scaling techniques on the power demand side, and (2) energy sources tuning mechanisms on the power supply side. Among those, the performance scaling techniques primarily leverage server power state tuning (e.g., DVFS and ACPI techniques [24], [25], [26]) and workload scheduling for load balance to accommodate runtime power budget or track the fluctuated renewable energy budget [11], [14], [15], [16], [27]. These approaches can forcefully cap power mismatches at the cost of performance degradation. Recently, a new tuning knob on the power supply side, the energy storage devices (short for ESDs, e.g., UPS batteries),

- *L. Liu, H. Sun, J. Xin, and N. Zheng are with the School of Electrical and Information Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi 710049, China. E-mail: {liulongjun, hsun, jxin, nnzheng}@xjtu.edu.cn.*
- *C. Li is with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Xuhui, Shanghai 200000, China. E-mail: lichao@cs.sjtu.edu.cn.*
- *T. Li is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611 USA. E-mail: taoli@ece.ufl.edu.*

Manuscript received 20 Dec. 2017; revised 13 Nov. 2018; accepted 2 Dec. 2018. Date of publication 12 Dec. 2018; date of current version 8 Sept. 2021. (Corresponding author: Hongbin Sun.)
Recommended for acceptance by W. Shi and C. Jiang.
Digital Object Identifier no. 10.1109/TSUSC.2018.2886382

is repurposed to shave peak power mismatching [6], [7], [8], [28], [29], [30], [31], [32], [33]. Compared with performance scaling schemes, the energy buffer technique can effectively improve power availability of datacenter while mitigating performance penalty.

We notice that UPS batteries manifest several disadvantages if they are used to improve power dependability and address the power mismatching issues: (1) batteries have limited lifetime cycle (approximate 2000 to 3000 cycles [34]). Frequent discharging can lead to a much shorter lifetime [35]; (2) large discharge current may lead to less usable capacity (known as the Peukert's law effect) [36]; and (3) to avoid battery overheating during charging, batteries cannot be recharged very fast with large charging current. In addition, the low energy efficiency is another major drawback of batteries – the, round trip energy loss can reach to 15-20 percent [37]. Therefore, *can we find a new way to gracefully handle the power mismatching on the power supply side while avoiding these limitations of batteries?*

In this paper, we propose a different power provisioning scheme—Hybrid and Hierarchical Energy Buffering (HHEB), which fully utilizes the advantages of incorporating various energy buffer technologies and deploying the energy buffers in different hierarchies of datacenter to efficiently handle power mismatching. The design can effectively improve power dependability and energy efficiency of datacenters. Specifically, we integrate super-capacitors (a.k.a., ultra-capacitors) with conventional UPS systems as a hybrid energy buffer to provide an additional layer of safety for datacenters in the event of unexpected power mismatches. Super-capacitors (SCs) have emerged as a promising alternative to batteries [38]. They have the following advantages: (1) high efficiency and low round-trip energy loss, (2) allowing fast charging and discharging with a high current, and (3) two to three orders of magnitude more life cycles than batteries [37], [38]. However, currently SCs are still too expensive for the large-scale, exclusive deployment in datacenters. As a result, the hybrid energy buffering system provide a more feasible and attractive solution.

When transmitting from homogeneous energy buffer to hybrid and hierarchical energy buffer technologies, challenges arise as the latter requires more intelligent power management schemes among various energy buffers in different hierarchies to achieve efficiency and economy: (1) for a given peak power mismatching scenario, there exists an optimal schedule of discharging that could provide the longest discharging duration. Note that the optimal discharging point often shifts as the available stored energy changes in either batteries or SCs, (2) for a given valley power charging opportunity, the energy buffers should be rapidly charged so that they can supply enough energy prior to the following peak power mismatching, and (3) the ESDs in different hierarchies should be coordinately managed for high efficiency and low cost. What is more, from the perspective of energy efficiency, the ideal usage pattern of heterogeneous energy buffers also depends on different power mismatching scenarios.

In this paper, we makes the following contributions:

- We explore hybrid and hierarchical energy buffers as new tuning knob on the datacenter power supply side to handle the irregular power mismatches. By comparing the cost, lifecycle, discharging rate, energy efficiency etc. among various ESDs, we demonstrate the design feasibility of heterogeneous and hierarchical energy storage buffering in datacenters.
- We propose HHEB, an efficient hybrid and hierarchical energy buffering based power provisioning architecture that could enable various energy storage devices to be efficiently integrated into existing datacenters for improving the dependability and efficiency. The architecture of HHEB is based on distributed and reconfigurable energy storage scheme which is easy to scale out and configure in datacenters.
- We present a hybrid and hierarchical power management framework, which can intelligently assign different ratio of the server loads to appropriate hybrid energy buffers in different hierarchies for high power dependability and efficiency during power mismatching events. The power management framework can autotune the load assignment and self-optimize its assignment performance.

The rest of this paper is organized as follows. Section 2 discusses related work; Section 3 provides the background and motivation of handling power mismatching. Section 4 characterizes the hybrid energy buffers and highlights the key design considerations. Section 5 presents the hybrid and hierarchical power provisioning architecture. Section 6 proposes the power management policies for HHEB. Section 7 describes our prototype system and experimental methodology. Section 8 presents the evaluation results and Section 9 concludes this paper.

2 Related Work

2.1 Novel Datacenter Power Provisioning Schemes

With the increasing of scale and capacity, modern datacenters become more power-constrained and carbon-constrained. To address these issues, many novel power provisioning schemes begin to spring up recently [2], [3], [4], [5], [6], [7], [8], [13], [14], [15], [16], [17], [18], [30], [39], [40], [41], [42], [43]. Wang et al. [30] proposed to virtualized power provisioning scheme in datacenters, their *vPower* can significantly improve system utilization and application performance when working in under-provisioned power infrastructure. Pelley et al. [5] presented a dynamic power provisioning scheme for datacenters. Their *Power-Routing* exploits shuffled topologies to dynamically connect the servers and diverse PDUs while balancing the workload across the PDUs for reducing the power infrastructure provisioning cost. Meanwhile, there are many renewable power provisioning schemes for datacenters to reduce carbon emission [13], [14], [15], [16], [17], [18]. In this paper, we propose hybrid and hierarchical energy buffer provisioning scheme for datacenters. Especially, we focus on dispatching hybrid energy buffering to dynamically and efficiently handle the power mismatching in the emerging power under-provisioned datacenters and renewable energy powered green datacenters.

2.2 Energy Dependability Study and ESDs in Datacenter

Recent efforts have also started to repurpose energy storage devices [6], [7], [8], [29], [30], [31], [32], [44], [45], [46], [47], [48], [49], [50], [51], [52], [53] to address datacenter energy availability issue for high power dependability while guaranteeing load performance.

Govindan et al. [6] discussed the benefits and limitations of leveraging energy storage device (ESD, e.g., lead-acid batteries) in datacenters to reduce datacenter peak power cost. Nonetheless, the proposed centralized architecture may incur 10-15 percent energy loss due to double-conversions. Kontorinis et al. [29] proposed distributed energy storage system (per-server UPS) to store energy during low load activity periods and use the energy to shave each server's peak. Both Yu [45] and Guo [46] proposed the battery based load scheduling and energy management policies in datacenter for cost minimization.

As battery manifests many limitations in energy efficiency, charging/discharging, etc., several recent studies have tried to explore other ESDs as new tuning knob [31], [32], [33], [47], [48], [49], [50]. Zheng et al. [31] exploited centralized thermal energy storage (TES) to shaving peak power in datacenters. As limited by the response time, they also combine the conventional UPS system to handle the frequent and transient peaks. Guo et al. [47] leveraged thermal storage to facilitate green energy integration and reduce the cost of brown energy usage. Li et al. [48] further studied fuel cell as energy storage device to efficiently handling power surges. Likewise, SCs have also grabbed certain attention in recent characterization work [32], [49]. Wang et al. [32] studied the multiple ESDs provisioning technology and placement options for datacenters. However, the major distinction of the two work includes two aspects: (1) ESD energy dispatching and controlling algorithms. Wang's work presents a theoretical framework for capturing important characteristics of different ESD technologies, and explore the trade-offs of placing ESDs at different power hierarchies. Their work emphasizes the characterization of multiple ESDs. However, we further propose the ESD energy dispatching algorithms for different workload power demands at the three power hierarchies. (2) ESD power control architecture and real prototype system. Wang's work investigated the energy storage provisioning - what, where and how much—in the datacenter for Demand Response (DR), and they propose a theoretical model about on energy storage technologies/characteristics. However, our work further study how to design an ESD power control and placement architecture at the three power hierarchies in datacenter, and how to implement the energy efficient hybrid and hierarchical design in a real prototype.

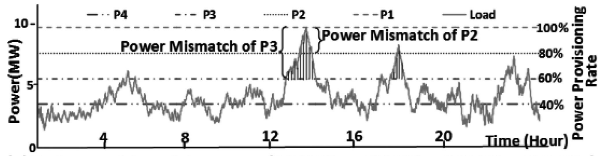
3 Background and Motivation

3.1 Power Under-Provisioned Datacenters

Conventional datacenter power infrastructures are commonly over-provisioned based on the nameplate rating power of all the servers, but this incurs significant power overhead and low power infrastructure utilization [2]. To this end, many datacenters today start to underprovision power infrastructures [2], [3], [4], [5], [6], [7], [8]. To detail the benefits and disadvantages of the schemes, we analyze the different power provisioning rates based on a

Google cluster workload trace [2], [32], as shown in Fig. 1a. We assume four different power provision rates (P1-P4). Among those, P1 is an over-provisioning scheme and can cover all peak demands. P4 is an under-provisioning scheme and only supplies 40 percent power budget for the datacenter loads. We define the maximum efficiency of power provisioning utilization (*MEPPU*) as:

$$MEPPU = \frac{\sum_{t=0}^{T} \Delta t\{P > P'\}}{T} \times 100\%. \quad (1)$$



(a)Under-provisioned datacenter($MEPPU_{P1}$ <$MEPPU_{P4}$, $CAP\text{-}EX_{P1}$>$CAP\text{-}EX_{P4}$)

Fig. 1. Modern datacenters in under-provisioned power infrastructure.

In Eq. (1),t is the accumulated time, during which power demands (P) exceeds the provisioned power budget (P'), T is the total load running time. Aggressively under-provisioning power infrastructure can yield high *MEPPU* and low infrastructure capital cost (capital cost is propotional to the provisioned IT power facility, estimated as $10-20 per Watt [3], [6], [8]). Nevertheless, the under-provisioning power infrastructure incurs more power mismatching, which degrade load performance if they are forcedly capped. Therefore, the power mismatchings are more easily appear in the power under-provisioning datacenter.

3.2 Renewable Energy Powered Datacenters

Provisioning clean renewable energy into datacenters can alleviate their carbon emissions. However, due to the intrinsic output fluctuation of renewable energy, intermittent power mismatching is one of the greatest challenges for the dependability of the renewable energy powered datacenters. Recent proposals leverage load deferment and load scheduling [9], [10], [11], [12] to match demand to the intermittent supply, which may violate the service level agreement and are not suitable for performance oriented datacenters. Another approach is to utilize largescale battery farms to shave the power mismatches for performance consideration.

During the peak power, the load can draw additional energy from batteries, and during the valley power, the surplus renewable energy can recharge batteries. Since the renewable energy generation is time-varying, it is critical for batteries to make the most of the opportunities of each power valley to store more energy. Therefore, the efficiency of renewable energy utilization (*REU*) is a crucial consideration to maximally utilize the green energy for intermittent power mismatches handling. The *REU* can be defined as:

$$REU = \frac{\left(\sum_{t=0}^{T} Bat_{RE} + \sum_{t=0}^{T} Load_{RE}\right)}{\sum_{t=0}^{T} Source_{RE}} \times 100\%. \quad (2)$$

In Eq. (2), Bat_{RE} is the renewable energy stored in batteries, $Load_{RE}$ is the renewable energy used for load and $Source_{RE}$ is the total amount of renewable energy generation.

Typically batteries have the upper bound of charging current and cannot timely absorb all the renewable energy during the very deep power valleys, which wastes renewable energy and leads to low efficiency of *REU*. Consequently, we need alternative energy storage devices without the limitation of upper-bound charging current, which can take advantage of the deep valley power mismatching opportunities to maximally absorb intermittent renewable powered and improve the power efficiency and availability.

4 Hybrid Energy Buffers: Characterization and Key Design Discussion

4.1 Characterization and Comparison

Hybrid ESD Design Space Discussion. Continuing technological advances have provided us increasing opportunities to

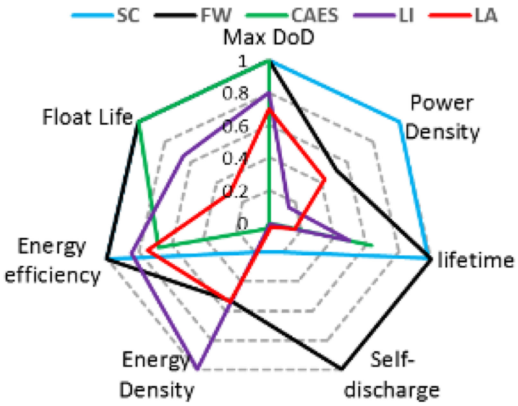


Fig. 2. Comparison of different ESDs.

employ competitive ESDs for improving the dependability and efficiency of emerging datacenter. In Fig 2, the radar chart compares different ESD technologies in terms of their key parameters (the results are normalized to the maximum value among the five ESDs). All the reported parameters are excerpted from [6], [8], [31], [32], [33], [34], [35], [36], [37], [38]. Power Density/Energy Density (Wh/L) determines the "volume" (real-estate) that needs to be provisioned in the datacenter to sustain the power demands. Lifetime (#discharging cycle) is the total number of charging/discharge cycles before the battery maximum charging capacity drops below 80 percent of the original battery capacity. Self-discharge (%) is the loss energy after charging even when they are not being charged. Float life (#discharging cycle) is the total cycle when battery stay at floating discharge/charge condition. Max DoD (%) is the maximum depth of discharge operations.

As can be seen from the chart, there are sharp differences among various ESDs. Super-capacitors (SC) have high power density, long lifetime cycle, high energy efficiency and depth of discharge (DoD). These advantages allow SCs to deliver large transient power. However, the SC has significant capital expense as mentioned above. Most characteristics of Flywheels (FW) are similar to SC, but it has the highest self-discharge rate. Therefore, FW is usually used for managing the brief interlude between the outage and generator start-up. Compressed air-based energy storage (CAES), has a relatively high specific energy but a low specific power, implying that it is better suited for holding a large amount of energy as long as this energy does not need to be discharged very fast. Lithium batteries (LI) has a high energy density, which facilitates it can provide the same energy capacity in a small volume. With the low cost and mature technology, Lead-acid batteries (LA) are the most commonly used ESD in datacenters, but many power and energy related characterizations are relatively low as show the red rectangle in Fig 5. Therefore, these ESDs offer very different trade-offs among the characterizations, which suggests that employing a combination of ESDs may allow higher efficiency

Taking batteries and SCs as a case of hybrid ESD, we first build up an energy storage characterization test-bed which consists a group of SCs (Maxwell 16V, 600F [54]) and LA-batteries (12V, 4AH). The batteries are connected to multiple servers by different relays. All the batteries are shared by the servers. We choose a group of datacenters workloads from Hibench (Table 1, Section 7) to evaluate the characterization of energy storage device.

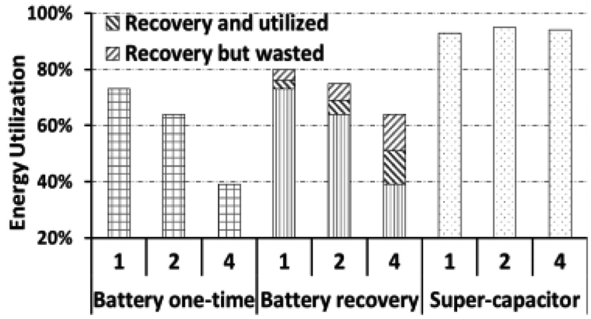


Fig. 3. Energy efficiency (1, 2, and 4 is the number of servers).

Energy Efficiency Analysis. One of the primary reasons for using SCs to buffer energy is that they incur negligible round-trip energy loss [37]. The energy efficiency of energy storage device can be calculated as below:

$$\eta ESD = \frac{\sum ENERGY_{discharge}}{\sum ENERGY_{charge}} \times 100\%. \qquad (3)$$

In Eq. (3), $\sum ENERGY_{discharge}$ is the total energy that is discharged from battery, $\sum ENERGY_{charge}$ is the total energy that is charged to battery. Our experimental measurements indicate that SCs can achieve 90-95 percent round-trip energy efficiency, as shown in Fig. 3. In contrast, lead-acid batteries have less than 80 percent efficiency even in the best case in our experiments. The energy efficiency calculation is based on detailed battery charging and discharging logs collected from our test-bed with different server power demands.

In fact, the efficiency of batteries can be even worse, depending on their usage patterns. There is a so-called recovery effect: batteries cannot release all of their stored energy in a one-time, high-current discharging–part of the stored energy seems to be "lost". During periods of no or very low discharge (by throttling server power demand), they can recover the energy "lost" to a certain extent [55]. Fig. 3 shows our characterization of different discharging scenarios with one, two and four servers, which reflect different power demands and battery discharging currents. The one-time discharging efficiency of the lead-acid battery decreases as we add more servers (i.e., increase the power demand). Given additional discharge cycles and enough recovery time, the battery efficiency can increase significantly (i.e., by 6~24 percent). However, this does not mean that one should always cap load power demand and wait for the battery to recover. This is because the energy waste due to server on/off cycles can be significant (i.e., account for nearly half of the recovered energy). Therefore, to improve energy efficiency, it is wise to use SCs to deal with power mismatching.

Charging and Discharging Comparison. Batteries and SCs manifest completely different charging/discharging features as battery stores energy electrochemically while there is no chemical reaction in SCs. SCs can be charged very fast without the limitation of upper-bound charging current, but neither does battery. We compare different discharging scenarios of batteries and SCs with different numbers of servers (Fig. 4). Our results show that the SC discharging voltage shows linearly declining trend irrespective of power demands. However, batteries exhibit a sharp voltage drop in light of large power demands since the chemical reaction process in batteries is slow and cannot release more power with a short time period. When handling power

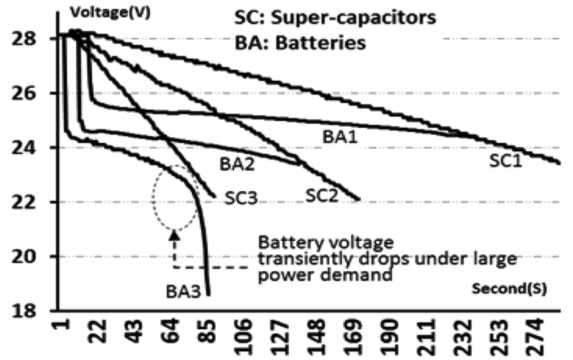


Fig. 4. Comparison of SC and battery discharging.

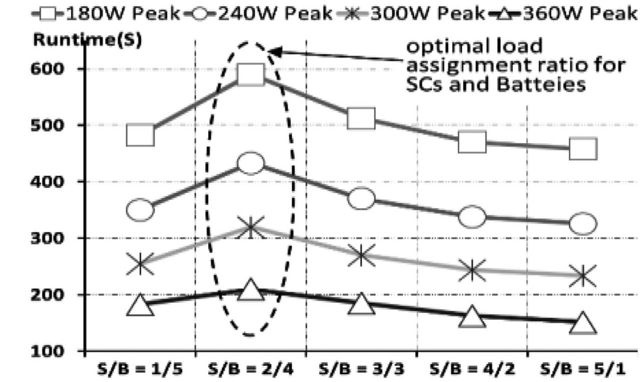


Fig. 5. Discharge duration comparison (S/B = m/n means m servers powered by SCs and n servers powered by batteries).

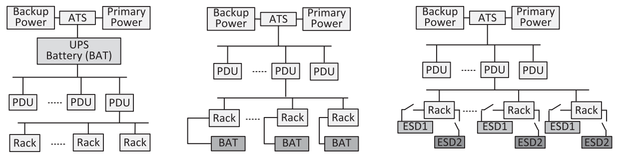
mismatching, the large peak power demands may cause battery voltage to transiently drop, which poses serious threat to server uptime. Therefore, it is important to avert using batteries to handle the large peak power mismatches. On the contrary, with the linear discharging properties, SCs are more stable and controllable for those scenarios.

4.2 Implications and Key Design Concerns

Based on the characterization above, it is obvious that no single type of ESD can provide a one-size-fits-all solution, but hybrid energy buffer has the opportunity to make the best use of their merits and overcome the limitations. In this paper, we leverage SCs and lead-acid batteries to constitute hybrid energy buffer for improving the dependability and efficiency of green datacenters. SCs can deliver high discharge current needed for dealing with large power mismatching while being recharged quickly between the events with high energy efficiency. However, the current cost of SCs is still high for large-scale deployment in datacenters. Inexpensive and conventional lead-acid batteries are deployed for handling large and mild peak power. Other ESDs that have similar characteristic to SCs, such as Lithium batteries, can also be exploited for the hybrid energy buffer. Nevertheless, in this study, we emphasize the scheme of dependable and efficient power system using hybrid and hierarchical energy buffers for emerging datacenters.

According to the hybrid energy buffer, how to effectively allocate different ESDs to handle various power mismaching is still a significant challenge. With our test-bed, we further perform experiments to explore how to jointly utilize hybrid ESDs to power servers. We first vary the number of servers assigned to the batteries and SCs to measure the maximum server runtime with constant power demands. In the experiments, whenever one energy storage device is depleted, the other will take over the entire load immediately via power switches. As Fig. 5 shows, Y-axis is the maximum server runtime, X-axis is the ratio of SCs and Batteries. There is an optimal load assignment that can provide the longest discharging time. It is clear that one should not heavily rely on either SCs or batteries. For example, by assigning heavy load on SCs, the server cluster runtime (uptime) can be decreased by 25 percent on average. Therefore, we should identify an optimal ratio to assign servers that powered by batteries or SCs for maximizing the server runtime.

The challenge of such load power assignment is that there is not a fixed optimal operating point under different total capacity of heterogeneous energy buffers and the time-vary shape of power peaks. The optimal server assignment actually depends on the current capacity of the heterogeneous energy buffers and the time-vary shape of power peaks. Therefore, we should dynamically identify the optimal operation point to distribute appropriate loads among different energy storage devices upon a power mismatching event.



(a) Centralized energy storage architecture (b) Distributed energy storage architecture (c) Heterogeneous energy storage architecture

Fig. 6. The different energy storage system architectures in datacenters.

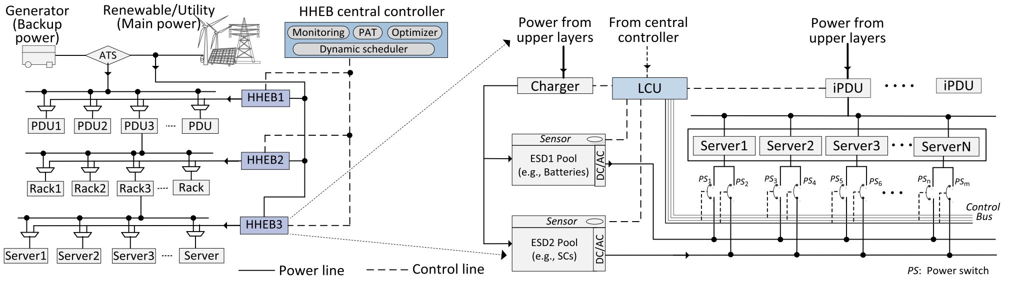
5 HHEB Architecture Analysis and Design

In order to effectively utilize various energy storage devices, we first study the ESD provisioning architecture in datacenters. We then analyze the pros and cons of leveraging current energy storage system to handle power mismatching. At last, we propose our HHEB power provisioning architecture in detail.

5.1 Current Energy Storage Architecture Analysis

Many research efforts focus on optimizing ESD power delivery topology to enhance datacenter power efficiency. The centralized (Fig. 6a) and distributed (Fig. 6b) topologies are two primary energy storage architectures in datacenters currently. In a centralized battery energy storage system, the UPS battery system locates on the critical path between the Automatic Transfer Switch (ATS) and the Power Distribution Units (PDU). When used to deal with the peak power mismatching (similar to [8]), it can only provide load shifting for the entire datacenter but cannot handle the peak shaving in a fine-grained manner. Moreover, the centralized UPS system commonly works online and always performs double converting (AC-DC-AC), which leads to 4-10 percent power losses [29].

Currently, IT giants such as Google, Microsoft and Facebook have explored the distributed power topology (Fig. 6b) in their datacenters. For instance, Facebook employs a cabinet of batteries for every 6 racks, or a total of 180 servers [56]. Their design is scalable in rack level and allows datacenters to shave peak power by using a fraction of the installed batteries. To avoid power double converting, it needs customized servers that support DC power. Google



(a) An overview of HHEB architecture in datacenter. (b) Zoom in view of HHEB architecture in server hierarchy

Fig. 7. (a) An overview of HHEB architecture in datacenter. (b) Zoom in view of HHEB architecture in server hierarchy.

mounts a battery in each server after Power Supply Unit (PSU) [57]. This design can completely avoid the battery double converting energy loss when shaving the peak power mismatching [29]. However, each server is assigned to a dedicated battery and multiple servers cannot share battery energy with each other to assist peak shaving. Furthermore, as the batteries are deployed in the inner chassis of servers, they are constrained by limited capacity. Note that both of the existing designs for datacenters are exclusively based on the homogeneous batteries and inevitably suffer the drawbacks of battery. Fig. 6c depicts our hybrid energy buffer topology, which provides opportunities to employ the pros and evade the cons of batteries and SCs when handling the power mismatching. The power switch based control enables datacenters to dynamically determine the distribution of server power demands between batteries and SCs. The batteries will offer bulk energy to the load since they can deliver large amount of energy slowly over a longer period of time while the SC pool will handle the transient peak power mismatching since they can be charged and discharged quickly.

5.2 HHEB Power Architecture Design

To improve the energy efficiency and dependability of datacenter power system, we propose the HHEB power provisioning architecture, as shown in Fig. 7a. The hybrid energy buffers are deployed in three layers of datacenter power distribution hierarchy. HHEB1 is deployed in the PDU hierarchy. HHEB2 and HHEB3 are respectively deployed at the rack and server level. The renewable/utility power charges ESDs of different HHEB hierarchies when the load power demands are lower than the provisioned main power budget.

The HHEB central controller is a key decision-making component that controls the power allocation of the three hierarchies. Fig. 7b shows the zoom-in view of HHEB architecture in server hierarchy. It includes two groups of ESDs (batteries and SCs). The power switches (*PS*) route energy from ESDs to different servers. The voltage & current of ESDs collected from sensors are transmitted to the Local Control Unit (LCU). The *PS* states (i.e., ON/OFF) as well as all the server power demands information (measured by the Intelligent Power Distribution Unit or IPDU) are transferred to the LCU too. LCU sends the above state feedbacks to HHEB central controller that makes operation decisions and sends command signals to LCU for controlling each power switch to distribute different energy sources for each server. Both the ESD architectures of the rack and PDU hierarchy are the same with the server hierarchy above. The energy allocation of the three hierarchies is coordinately managed by the HHEB central controller. In our current implementation, the HHEB central controller is a low-power server that hosts our power management algorithms, such as dynamic scheduler, optimizer, etc. It contains various power management APIs, e.g., *power allocate* (BATTERY *n*, *SC m*, SERVERS *k*) controls hybrid energy allocation for different servers; *switching* (RELAY *i*) turns the power switch (*PS*) ON/OFF; *charging* (BATTERY *n*, *SC m*) and *discharging* (BATTERY *n*, *SC m*) control the energy transferring; and *SoC* (BATTERY *n*, *SC m*) monitor the power state of ESDs; *DVFS* (SERVERS *k*) tunes the server power demands. These APIs provide basic functions for the HHEB power management framework.

Compared with the flat and single-tier energy buffering architecture, HHEB manifests two significant advantages: (1) According to the various power mismatching characterizations in different datacenter power hierarchy, we can deploy different ESDs in each HHEB hierarchy to improve the efficiency of peak power handling. For example, using energy storage devices such as batteries and supercapacitors to make up for short duration power shortfalls caused by the power surges in server hierarchy. Fuel cell or CAES can be deployed in the rack or datacenter hierarchy for handling the slow and longer power mismatching. Note, the Fuel and CAES deployed different hierarchy of HHEB will be studied in our future work. In this work, we focus on the batteries and SCs. (2) the relays in different hierarchies can quickly assign the energy pool to different load, which can avoid workload performance degradation cause by workload/VM migration or balance.

6 HHEB POWER MANAGEMENT FRAMEWORK

In this section, we present power management framework of HHEB, which primarily integrates power mismatching prediction, dynamic load assigning, energy buffer optimization and hierarchical energy buffer management policy.

6.1 Problem Formulation

At the beginning of each control interval, the controller obtains the current available capacity of ESDs (ΔSC and ΔBA) based on the feedback information from sensors (Note that we here take batteries (ΔBA) and SCs (ΔSC) as a case of hybrid ESDs). We assume the total power mismatching during the control interval is ΔPM. We define R_λ as the ratio of servers powered by SCs, therefore, the number of server powered by SCs is $NumS * R_\lambda$, where $NumS$ is the total number of servers. Likewise, the number of server powered by batteries is $NumS * (1 - R_\lambda)$. The controller assigns the energy buffer based on the above four variables (ΔBA,ΔSC, ΔPM, and R_λ) to handle the peak power mismatching events. The energy efficiency (*EE*) and server

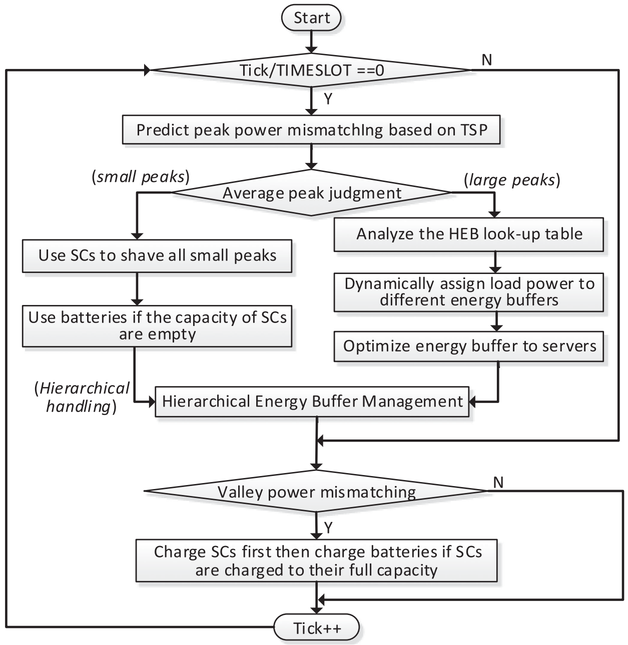


Fig. 8. An overview of HHEB power management framework (includes prediction, small/large peak handling, valley power charging, and hierarchical energy buffer management).

downtime (SD) at the end of each control interval (t) can be calculated based on the four variables:

$$f_t : (\Delta BA, \Delta SC, \Delta PM, R_\lambda) \rightarrow (SD, EE) \tag{4}$$

$$\text{SD} = \sum_{t=0}^{T} Du(\text{t})\{[(\Delta BA + \Delta SC) - \Delta PM] \ < \ 0\} \tag{4-1}$$

$$\text{EE} = \frac{\Delta BA_{discharging} + \Delta SC_{discharging}}{\Delta BA_{charging} + \Delta SC_{charging}} \times 100\%. \tag{4-2}$$

In Eq. (4-1), *SD* (*Server Downtime*) is calculated by the aggregated duration ($Du_{(t)}$) which reflects the total stored energy in batteries (ΔBA) and supercapacitors (ΔSC) is less than the energy of power mismatching (ΔPM). In Eq. (4-2), *EE* is the overall energy efficiency of hybrid energy buffers which is calculated by aggregated discharging and charging energy from battery (ΔBA) and supercapacitors (ΔSC). As the values of ΔBA and ΔSC are not fixed during each power mismatching period, our power management goal is to minimize the *SD* and maximize the *EE* by adjusting the ratio of energy allocation. The two optimization goals (*SD* and *EE*) are not conflicted but interactional. We will solve the problem by constructing a dynamic load scheduling and balanced policy, an optimized energy allocation method, and a hierarchical energy management scheme and continuously optimizing in the whole lifetime of energy storage device, which can efficiently minimize *SD* and maximize *EE* simultaneously.

6.2 Dynamic Load Scheduling for Hybrid ESDs

A key problem in the design space of HHEB is how to assign heterogeneous buffers to the most appropriate load. If we fixed the heterogeneous energy pool and perform load balance, the load migration may easily cause workload performance degradation. HHEB dynamically distributes batteries and SCs to shave various power peaks in each timeslot (Similarly to general power control time-slot, we set the default value of time-slot is 10 minutes. Longer time-slot duration may aggravate the effectiveness of prediction error). Fig. 8 shows the power management framework.

Prediction. To identify the average peak power char-acterization (e.g., small peaks or large peaks) of next time-slot, we employ time series prediction (TSP) method [58] to predict the peaks of each hierarchy. Specially, we leverage the classical triple exponential prediction (Holt-Winters exponential prediction) algorithm [59] to periodically predict the power demands, which can analyze the nature of the history and current data, extract meaningful statistics trend and predict future values. The algorithm maintains two groups of series data: the peak power and valley power. It predicts the peak power demands (P_{peak}) and valley power (P_{valley}) of next time-slot. The difference of P_{peak} and P_{valley} ($\Delta PM = P_{peak} - P_{valley}$) is the net amount power that needed from the energy buffers. Note that we select a time series prediction method that is effective for the datacenter power consumption patterns, but any sophisticated prediction approaches can be integrated into our power management framework. Since the prediction interval is 10 minutes, the power value in different timeslot also means the energy demand information.

Small Peaks Handling. When the average height of predicted power mismatching is mild and the duration is short, the power management controller treats the batteries and SCs as a two tier energy storage system. Either batteries or SCs can handle these small peak power mismatches. In order to enhance energy efficiency, the controller preferentially assigns all loads power on SCs ($R_\lambda = 1$). This is because SCs have much better roundtrip energy efficiency and they can be swiftly charged and discharged without degradation. Only when all the SCs are used up, the controller will turn on all the battery relays and assign all server loads on batteries ($R_\lambda = 0$) to compensate the energy shortages. In brief, SCs are aggressively used to handle the small peak power mismatching for high energy efficiency while maintaining minimal server down time by employing batteries as supplement during the interval when SCs are used up.

Large Peaks Handling. The power management controller treats batteries and SCs as a unified energy buffer when the predicted average peak power mismatching is significant and the duration is long (large peaks). In other words, the controller schedules all the loads on batteries and SCs simultaneously to jointly shave peaks.

To maximize energy efficiency and minimize server downtime, we should carefully allocate an optimal R_λ ($0 < R_\lambda < 1$). To this end, the power management controller maintains a power allocation table (PAT) for its hybrid energy buffers. This table specifies initial and coarse grained load assignments on batteries and SCs. Each entry of the power allocation table contains the available energy levels of the battery and SC pools, power demands and the server ratio that indicates the fractional servers powered by SCs and batteries. The initial value of each entry is obtained via profiling in a pilot scheme like Fig. 6. The profiling values in the table are not fixed all the time, and they can be optimized and updated (detailed in Section 6.3). *Algorithm 1* shows the pseudo code of the server loads assignment (Lines 1-11). Based on the available energy buffer and predicted average power mismatching value at each time-slot, the controller can find the energy allocation ratio R_λ or similar R_λ in the PAT and dynamically control the on/off power switches to assign different ratio servers powered by SCs or batteries. However, as it cannot profile all scenarios of available energy buffer and power demands, the number of entries in PAT is limited. Therefore, it may be difficult to find an optimized energy allocation ratio R_λ in such initial PAT.

Algorithm 1. Large Peak Power Mismatching Handling and Energy Dispatch

Input: Current SC capability: $SC_{initial}$, Batter capacity: $BA_{initial}$ and predicted power mismatching $\Delta PM(\Delta PM = P_{peak} - P_{valley})$;
Output: The energy allocation scheme and ratio R_λ

1. Obtain current SC capacity: $SC_{initial}$, Battery capacity: $BA_{initial}$, and predicted power mismatching ΔPM $(\Delta PM = P_{peak} - P_{valley})$;
2. **For** table *index* = 1 to n // *search the look-up table PAT*
3. **If** $(SC_{index} == SC_{initial}$ && $BA_{index} == BA_{initial}$ && $P_{index} == \Delta PM)$
4. find_index = index;
5. **End**
6. **End**
7. **If** *(find_index == 0) //does not find a matched entry*
8. *find_index = Similar*($SC_{initial}$,$BA_{initial}$, ΔPM); *//search the most similar value*
9. **End**
10. Server ratio $R_\lambda = R_\lambda(\text{find\_index})$; *//Find the ratio in PAT;*
11. Allocate different numbers of servers to SC and BA based on R_λ;
12. Collect running results at the end of the time-slot.
13. **If** *(index == 0) //new entry (new energy buffer capacity & power demand)*
14. Round*($SC_{initial}$, $BA_{initial}$, P); //format data, P is the actual power demand*
15. *Add* $\{SC_{initial}, BA_{initial}, P, R_\lambda\}$ *to the PAT look-up table;*
16. **Else** *//update the existing entry of the PAT table*
17. **if** $(SC_{end}/BA_{end} > SC_{initial}/BA_{initial})$
18. $R_\lambda = R_\lambda + \Delta r$;*//SC receives increased server assignment*
19. **Else If** $(SC_{end}/BA_{end} < SC_{initial}/BA_{initial})$
20. $R_\lambda = R_\lambda + \Delta r$;*//BA receives increased server assignment*
21. **End**
22. Update $\{SC_{initial}, BA_{initial}, P, R_\lambda\}$ in the PAT look-up table;
23. **End**

6.3 Optimizing Energy Buffering Allocation

As mentioned above, the PAT table cannot always guarantee the optimal load assignment results because (1) the limited profiling data are based on a pilot run and can be less accurate, and (2) with the battery and SC aging, their ability of handling power mismatching will decline. Therefore, the table needs to be dynamically updated.

To ensure effectiveness, the controller updates the PAT table during runtime. *ALGORITHM 1* shows the pseudo code of the optimization operations (Lines 12-23): (1) adding new entries into the table, and (2) updating the existing entry. It first collects the running results at the end of the time slot, which includes the real power mismatching value and server load allocation ratio of current time slot. When adding a new entry, the results are formatted and become coarse grained to avoid too many entries in the table. When updating the existing entry, the controller checks the remaining capacity in SCs and batteries. If the actual battery capacity decline rate (Line 17) is faster than expected (e.g., due to internal wear-out, batteries were assigned too much load and have higher discharge rate than SCs), the controller will increase the load ratio by $\Delta r = 1\%$ (default value) to increase the usage of SCs in future allocation. If the actual battery discharging rate is slower (Line 19) than expected, HHEB system will reduce the load ratio to decrease the usage of SCs. This optimization operation is to balance the using of SCs and batteries for minimum server downtime.

6.4 Hierarchical Energy Buffer Management Algorithm

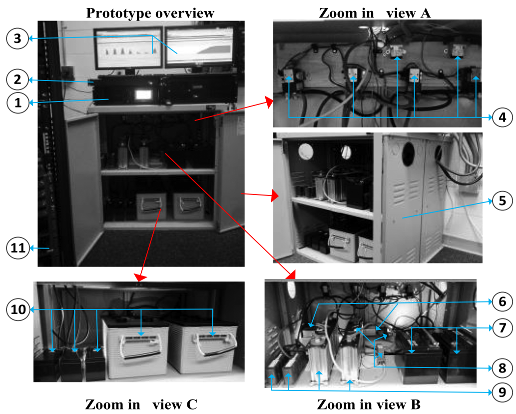
When there is insufficient energy in a single layer of ESD for power mismatching handling, we will leverage the hierarchical ESD energy management strategy to assist the power mismatching handling (Algorithm 2).

We first combine the three separated ESD hierarchies (server layer, rack layer and PDU layer) into two group of ESD units (Server-Rack unit and Rack-PDU unit). Each ESD unit includes a parent layer and a child layer. The ESD energy in parent layer can be distributed to handle the small/large peak power mismatching appeared in child layer. For the small peaks handling, we preferentially assign SC energy in parent layer to compensate the energy shortage of child layers for high energy efficiency. If the SCs are used up, the batteries energy of parent layers will be tapped into the child nodes. For the large power mismatching handling, the SCs and batteries in parent layer are jointly utilized to compensate the energy buffer in child layer for high energy availability. According to the energy allocation ratio (R_λ) in the child layer, SCs and batteries in the parent layer are simultaneously tapped into the child nodes for the large peak handling. When there is a slack in power usage, the hierarchical ESDs will be recharged in sequence. The SCs in child hierarchy are given higher priority to be recharged. Then, the batteries in the child layer will be recharged. If all the ESDs in the child layer have been recharged to their full capacity, the SCs and batteries in parent hierarchy will be recharged in turn.

According the *Algorithms* 1 and 2, we can see that SCs are preferentially employed for peak power shaving, which is because SCs have higher self-discharging current than batteries. With frequent charging/discharging operations, the self-discharge effect of SCs can be effectively mitigated.

7 Evaluation Methodology

We build a scaled-down prototype to evaluate our design and power management framework based on the proposed HHEB architecture. As shown in Fig. 9, the platform includes several small and large batteries/SCs connected by relays to power different servers. There are six two-way relays in our prototype which can simultaneously connect to six servers. The servers are mounted on the rack and respectively connected to IPDU. The IPDU can switch ON/OFF server power supply, report the server power draw every second and send it to the controller by SNMP commands over the Ethernet. Any power management algorithm can be integrated in the controller to monitor and control all components in our prototype. The platform also allows hierarchical deployment of various ESDs. With the rerouting cables and power switches, we can easily deploy batteries and SCs in rack or PDU



①: Control node (Siemens S7-200 Programmable Logic Controller)
②: Control (running power management APIs in OpenStack)
③: System real-time running state monitoring
④: Relays (two-way relays)
⑤: Heterogeneous energy storage cabinet
⑥: Inverter (two 1000W DC to AC power inverter)
⑦: Lead-acid batteries (24V DC system)
⑧: Sensors (voltage, current and temperature)
⑨: Small and large super-capacitors (SCs)
⑩: Small and large LA-batteries (24V DC system)
⑪: Server rack (low power servers)

Fig. 9. Full-system prototype of HHEB - a hybrid and hierarchical energy storage research platform.

hierarchy to fully evaluate the design of our hierarchical energy buffer technology (notes, the small batteries and small SCs are deployed in the PDU level, which is a 24V ESD system and their total capacity is 35Ah. The large batteries and SCs are deployed in the Rack level, which is a 24V ESD system and the total capacity is 200Ah). Our platform can be deployed in either conventional power under-provisioned datacenters or renewable energy powered datacenters to handle the power mismatching.

Algorithm 2. Energy Coordination in Different ESD Hierarchies

Input: The available capacity of battery and SC in different hierarchies of datacenter.
Output: The power supplied of ESD from the parent node.
1. Leveraging ESD nodes to handle small peaks in the child hierarchy
2. **If** ($\sum P_{ESD\_child\,\mathrm{mod}\,es} < P_{demand}$)//current ESD energy cannot cover all the server power demand
3. Using the SC_{parent} energy in the parent hierarchy to shave the small peak ;
4. **If** ($SC_{parent} < 0$) //if the SC energy is not enough for the peaks
5. Using the $Battery_{parent}$ in the parent hierarchy to shave the small peak;
6. **End**
7. Leveraging ESD nodes to handle large peaks in the child hierarchy
8. **If** ($\sum P_{ESD\_child\,\mathrm{mod}\,es} < P_{demand}$)//current ESD energy cannot cover all the server power demand
9. Jointly using the SC_{parent} and $Battery_{parent}$ *in the parent hierarchy to shave the large power peak ;*
10. $P_{SC} = R_{\lambda} * P_{demand}; P_{Battery} = (1 - R_{\lambda}) * Pdemand;$ *//eney allocation for battery and SC*
11. **End**

We choose various datacenter workloads from Hibench [60] and CloudSuite [61]. Hibench contains nine typical Hadoop workloads (including micro benchmarks, HDFS benchmarks, web search benchmarks, etc.). Cloud-Suite benchmarks are based on real-world software stacks and consist eight popular applications in today's data centers. As shown in Table 1, we select eight workloads from five classified categories. Within each experiment, a workload can be executed iteratively.

TABLE 1
The Evaluated Workloads [63], [64]

Workloads (Abbr.)	Category	Peak
Page Rank Algorithm of Mahout (**PR**)	Web Search Benchmarks	**Large Peaks**
Word Count Program on Hadoop (**WC**)	Micro Benchmarks	
Data Analysis (**DA**)	CloudSuite Benchmarks	
Web Search (**WS**)	CloudSuite Benchmarks	
Media Streaming (**MS**)	CloudSuite Benchmarks	**Small Peaks**
Dfsioe (**DFS**)	HDFS Benchmarks	
Hivebench (**HB**)	Data Analytics	
Terasort (**TS**)	Micro Benchmarks	

Our server system kernel can be configured with the on demand frequency scaling governor. We can set the low frequency as 1.3GHz and the high frequency as 1.8 GHz. To fully evaluate our peak power management policies, we divide the eight workloads into two groups, one group runs on the high frequency and the other group runs on the low frequency. In this way, we can construct two general peak shapes (small peaks and large peaks, which is defined based on different datacenter power demand) to fully evaluate our power management policies. Note that our method is similar to [8].

In our experiments, the controller can collect the utility power consumption of all the servers via IPDU. We set a maximum power drawn from utility budget, e.g., 260 W. Whenever the server power demands exceed 20 percent of the utility power budget (peak occurs), we treat the peaks as large peaks. Otherwise, when the overload demands are less than or equal to 20 percent, we treat the peaks as small peaks. The controller would tap into the energy stored in the energy buffers. Oppositely, the remaining energy can charge energy buffers when the server power demands are lower than the budget.

8 Experimental Evaluation

This section evaluates the benefits of provisioning hybrid and hierarchical energy buffers in datacenters. To be more specific, we evaluate the HHEB design in two steps. First, we compare the performance of hybrid energy buffer (HEB) with five kinds of power management schemes as summarized in Table 2. Second, we further evaluate the benefits of hierarchical energy buffering technology. As shown in Table 2, BaOnly is a representative peak power management technique similar to prior work [8], which only uses homogeneous Lead-acid batteries to shave peak power because lead-acid batteries account for over 97 percent of industry batteries [1] and they are widely deployed in datacenter due to their technical maturity, low cost, and easy maintenance. Note that with BaOnly, the servers are still mainly powered by utility grid when there is no peak power. Although BaFirst and SCFirst both use hybrid

TABLE 2
The Evaluated Power Management Schemes

Schemes	Architecture	Method description
BaOnly	Battery only	Only use battery to handle power mismatch
BaFirst	Hybrid (Battery+SC)	Discharge batteries first, then SCs if the capacity of batteries are empty
SCFirst	Hybrid	Discharge SCs first, then batteries
HEB-F	Hybrid	Load-aware assignment based on power demand value of the last time-slot
HEB-S	Hybrid	Load-aware assignment based on statics and limited profiling information
HEB-D	Hybrid	Load-aware assignment based on our dynamic and optimized power management framework

energy buffers, they lack intelligent server allocation policies and only employ a prioritybased method to handle power mismatches. The HEB-F and HEB-S are two naïve implementations of HEB. The HEB-F assigns the heterogeneous energy buffers to different servers based on the power demand information of last time-slot. The HEB-S assigns load power based on a static profiling table that has limited entries. The HEB-D is our proposed dynamic and optimized power management framework.

The purpose of comparing HEB-D with HEB-F and HEB-S is to understand the impact of reduced prediction error rate on performance improvement. To fairly compare the performance of battery only and hybrid energy buffers, their total capacity is set to the same by configuring the small and large SCs and batteries in the prototype (the initial ratio of SCs and batteries is 3:7). Note that this study mainly compares systems with equal storage capacity (so that they have the same worst-case emergency handling capabilities). The reason why we did not compare "equal size" or "equal cost" systems is that they are technology-/vendor- dependent. The capacity of SCs has a direct impact on the performance and lifetime of our systems. For "equal-cost" and "equal-size" designs, it is very hard to tell if the improvement is a result of our optimization scheme or a result of the capacity change due to different SC technologies.

8.1 Prototype System Running Profiling

We capture a piece of system running profiling trace from our prototype as shown in Fig. 10. The dashed line is the fluctuant solar power supply and the black line is the workload power demand. The dotted line shows the hybrid energy buffer allocation ratio R_λ. Y-axis (left) is the scale of power and the secondary Y-axis (right) indicates the scale of R_λ. X-axis is the control duration. Every point in the X-axis represents each time-slot. We can see that the ESD energy allocation ratio $R_\lambda$obviously varies with the variation of power mismatching. The larger value of $R_\lambda$means more SC energy is allocated to shaving the power mismatching. When the power demand is less than the power supply, the ESD will be recharged and the ratio will stop varying. Actually, the variation of $R_\lambda$in each time-slot is different because the available capacity of ESD is different and the self-optimization algorithms will adjust the ratio in each time-slot based on the ESD state.

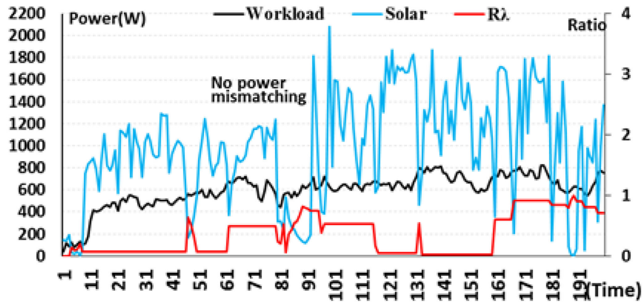


Fig. 10. Prototype system running profiling of HHEB.

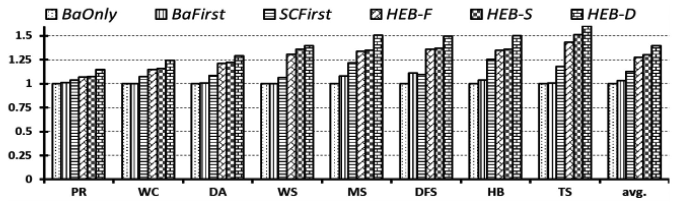


Fig. 11. Energy efficiency comparison of different power policies.

8.2 Efficiency Improvement: Energy Efficiency and REU

The energy efficiency includes the ESD discharging efficiency and renewable energy utilization. The energy efficiency is an important metric for the emerging ESD powered datacenters.

Energy Efficiency Comparison. To improve the efficiency of energy storage systems, one must carefully assign and utilize both SCs and batteries to obtain maximal energy efficiency.

Fig. 11 shows the overall energy efficiency measurement. Compared to a conventional battery-only power provisioning scheme, the heterogeneous energy buffers yield a visible efficiency improvement. The reason why *BaFirst* is very close to a battery only design is that *BaFirst* always charge/discharge battery first which reduces the chances of SCs utilization. If we always discharge the SC first, we can greatly reduce energy loss such as SCFirst, but when the SCs are depleted, batteries would have to handle all the high current drawn which still leads to efficiency degradation. Therefore, employing load-aware assignment to balance the usage of SCs and batteries can achieve better efficiency improvement (e.g., HEB). The energy assignment of HEB-F is based on the former power demand information which is a naïve prediction scheme and may lead to incorrect energy assignment. The errors in prediction decrease energy efficiency. The HEB-S often makes a suboptimal energy assignment as it only has a coarse-grained profiling table. In contrast, HEB-D can achieve better energy efficiency. In addition, HEB-D manifests higher efficiency on both small peak workloads (as SCs are preferentially used) and large peak workloads (as loads are dynamically allocated with energy between batteries and SCs) via our proposed policy.

Renewable Energy Utilization Comparison. We also present the benefit of hybrid energy buffer provisioning in light of renewable powered data centers. As mentioned in Section 3, it is critical to improve the renewable energy utilization (REU) for storing more green and clean energy to handle power mismatches in renewable datacenters. Compared with pure battery equipped systems, SCs can absorb renewable energy without upperbound of charging current, which can achieve more energy utilization. We tap into solar

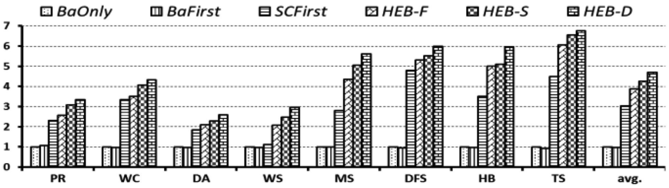


Fig. 12. Renewable energy utilization of different power policies.

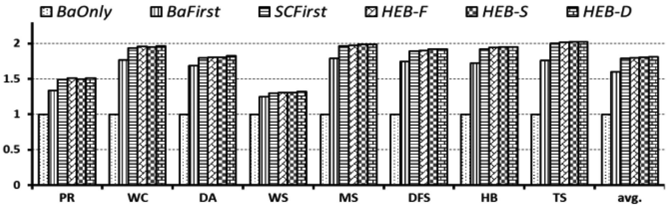


Fig. 13. Battery lifetime comparison of different power policies.

power to our prototype system instead of utility power to evaluate the REU. Note that we have a small solar generation system on the roof of our Lab, which can provide real solar power for our experiments. The results show that only if introducing SCs to the energy buffer, the REU can be significantly improved (e.g., BaFirst, SCfirst and HEB). As BaFirst gives the first priority to batteries, it may lose some chances to absorb renewable energy with large charging current. SCFirst and HEB always utilize SC first to absorb renewable energy; they have very similar REU (Fig. 12), and all of them improve the REU about 81 percent on average compared with the pure batteries provisioning scheme.

8.3 Dependability Comparison of Different ESD System

Battery Lifetime. Battery lifetime is related with the mean time to repair (MTTR) of battery in ESD system. The longer lifetime means drastic reductions in MTTR. One of the original intentions of introducing SCs as hybrid energy buffers is to protect batteries from large current discharging and prolong their lifetime. We use the Ah-Throughput Battery Lifetime Model [62] to present the anticipated battery lifetime based on detailed battery usage logs. As shown in Fig. 13, the SC preferential power management policy has more battery life cycle since batteries are used as backup (e.g., SCFirst and HEB). The HEB has better battery lifetime improvement as it only uses SCs to shave small peaks and jointly utilizes SCs and batteries to shave large peaks for protecting batteries from large current discharging. The HEB-D can improve the battery lifetime by 4.7X compared to the BaOnly scheme. Compared to the lifecycle of SCs, battery lifetime is the bottleneck of heterogeneous energy system lifespan. Longer battery lifetime implies lower replacement and maintenance cost of HEB.

Server Downtime. In our experiment, server down-time is the aggregated time during which server power demands exceed power budget but the energy buffers do not have enough power to shave the peak. We chose the least recently used servers to shut down when we have to. Note that in this paper we do not use other control knobs such as DVFS for directly comparing the dependability of different ESD power management schemes. In the real datacenter, one can put servers into low-power modes or perform VM migration to avoid peaks as well as the long restart-up time. Therefore, our evaluation of server

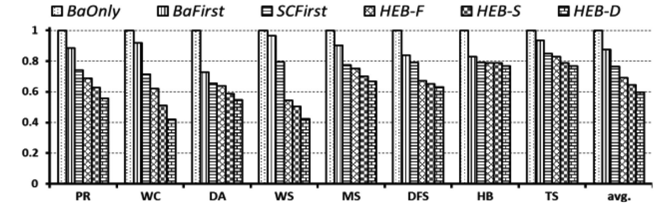


Fig. 14. Server downtime comparison of different power policies.

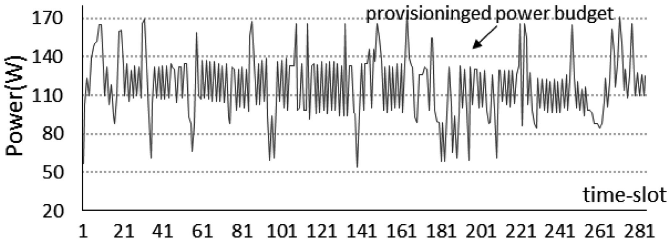


Fig. 15. Peak characteristics at server-hierarchy.

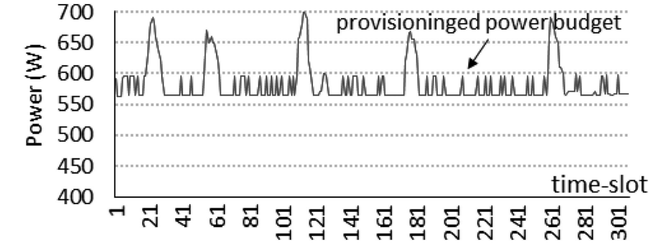


Fig. 16. Peak characteristics at rack-hierarchy.

downtime reflects the average availability yielded by a power management scheme. To compare the server downtime of different power management policies, we intentionally lower the utility power budget to trigger server downtime. Due to the Peukert law's effect, it is difficult to adopt BaOnly to handle the large peaks. Doing so may lead rapid drop of battery voltage, especially when the batteries have low SoC (State of Charge). The server downtime can be mitigated with integrated SCs in the ESDs (Fig. 14). As can be seen, HEB can always maintain the longest discharging duration by dynamically adjusting the server assignment between SCs and batteries. The HEB-D can reduce more server downtime (41 percent), especially for the large peak workloads.

The Availability of Hierarchical Energy Buffer. Based on our prototype system, we reconfigure the power relays to build a layered energy buffer system with SCs and batteries. We then evaluate the benefits of hierarchical energy buffering technology. According to the hierarchical energy dispatch strategies, we can dynamically allocate the ESD energy in different layers to handle various power mismatching.

Based on the statistics of workload power traces, we find that the power peaks in different hierarchies show different characterizations. The power peaks are small and dense at the server-hierarchy, as shown in Fig. 15. At the rack-hierarchy, the peaks are large and sparse, which because multiple power demands from server-layer may overlapped together at rack layer, as shown in Fig 16. Based on the different characterizations at different layers, we can provisioning corresponding ESDs to handle different peaks for high energy efficiency, we further can adaptively dispatch available energy to different layers for compensating energy lack.

We evaluate the energy compensation effects from SCs in parent hierarchy to batteries in child hierarchy as shown in

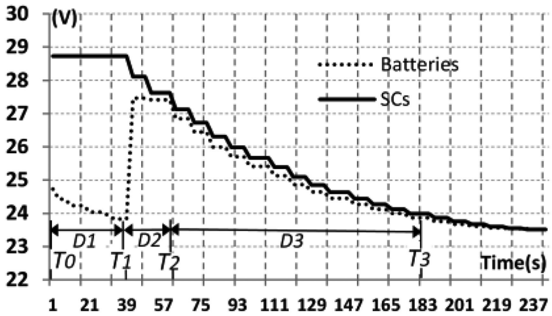


Fig. 17. Energy availability evaluation for the hierarchical energy dispatch strategies.

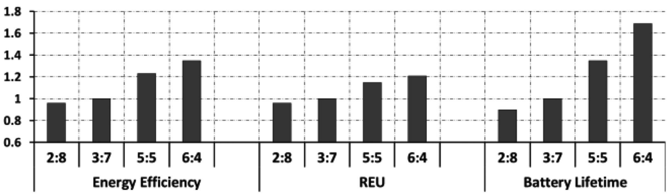


Fig. 18. The impact of different capacity ratios for efficiency and dependability (m:n is the capacity ratio of SCs & batteries. All the metrics are normalized to ratio of 3:7).

Fig. 17. At the beginning (time T0), only batteries are employed to shave peak power mismatching, which leads to battery voltage drops quickly from 24.9V to 23.8V. Then, the SCs in the parent hierarchy of HHEB are discharged for batteries charging during D2 (time T1 to T2), which evaluates the energy compensation effects from the ESD in parent hierarchy to the ESD in child hierarchy (Section 6.4). As the battery and SC are reached to the same voltage, they are discharged for load simultaneously when power peaks appear (D3: T2 to T3). After the duration D3, the voltage of batteries reached to 23.8V again, which is the same with the voltage of only utilizing batteries to shave peaks before. We can see that the battery availability is effectively extended by 3.2X (D3/D1) because of the energy compensation from the parent hierarchy. The improvement of availability is related to the capacity of ESD in different power hierarchies.

The benefit of hierarchical ESD technology can compensate the urgent energy requirement of different ESD hierarchies, while enabling flexible energy provisioning and adaptive peak power shaving. Moreover, compared with the single-tire placement of ESD, placing ESDs in multiple hierarchies can effectively improve the availability of ESD system.

8.4 ESD Capacity Variation Discussion

We evaluate the impact of different capacity provisioning for heterogeneous energy buffers. First, by keeping the constant total capacity of the energy buffer, we adjust the capacity ratio between SCs and batteries. In detail, we adjust the Depth-of-Discharge (DoD) of energy buffers to generate different capacity ratios for batteries and SCs. For example, given 8Ah battery, we set the targeted DoD level as 60 percent. Its useable capacity is 4.8Ah (8Ah*60 percent). Our controller can disable the utilization of batteries once it hits its DoD threshold. We iteratively run the eight workloads with HEB-D power scheme and respectively obtain the average performance of energy efficiency and battery lifetime, as shown in Fig. 18. The results show that the more ratios of SCs can obtain better performance improvement. Moreover, the impact of the capacity ratio is different across the four metrics. The battery lifetime has the most significant improvement as more SCs can be used to shave peaks. The improvement of energy efficiency and server downtime gradually becomes constant.

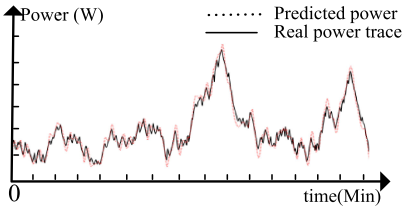


Fig. 19. A typical workload power trace prediction demonstration.

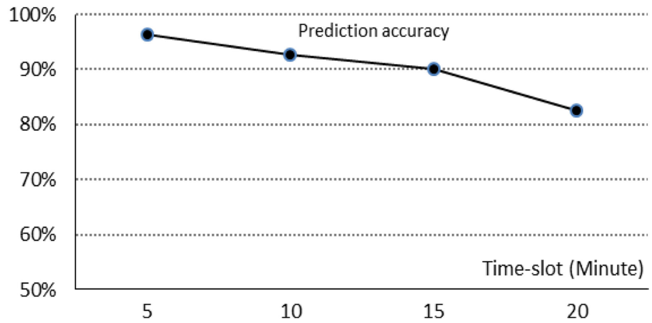


Fig. 20. Prediction accuracy variation with different time-slot.

8.5 Prediction Accuracy Analysis and Discussion

Peak prediction accuracy algorithms affect the effectiveness of energy dispatching for HHEB system. We evaluate various real-world data center traces based on different peak height (PH), peak width (PW), and peak frequency (PF). Based on the accuracy statistic results, we find that two key factors affect the accuracy of peak prediction algorithm: (1) *various peak features*. For the mild and wide peaks, the prediction accuracy of our algorithm can research to 97.8 percent. However, the prediction accuracy will reduce to 89.5 percent when it predicts the narrow and high peaks. For the high peaks, the prediction error may lead to severe energy shortage and controlling failure. However, the prediction error for the small peaks may not cause energy shortage if the stored energy in ESDs is enough to shave the prediction error. Fig. 19 show a typical power trace prediction with our time series prediction (TSP) method. The prediction accuracy can reach to 92.7percent. (2) *Different durations of time-slot*. General energy dispatch time-slot is 10 minutes. Many previous studies have employed a default time control periods to evaluate their energy dispatch design. We set the "10-min" as a default value of time-slot. Longer time-slot duration may aggravate the effectiveness of prediction error. For the short duration of time-slot, the prediction error in the timeslot may be alleviated if the prediction precision recovered in the following timeslot, but it may cause frequent energy dispatching operations and lead to unnecessary controlling overhead. Fig. 20 present the prediction accuracy variation with four different time-slot (5 Min, 10 Min, 15 Min and 20 Min).

In this study, we leverage the peak power prediction method to identify the small or large power peak in each power demand timeslot (e.g., 10 minutes by default). Based on the prediction result, we decide when to use the power

TABLE 3
The Cost Constitution of HHEB

SCs	Batteries	Inverter	Sensors	PLC	Relays	Cables
32.4%	23.1%	12%	13.5%	9.7%	4.9%	4.4%

TABLE 4
The ROI in Different Peak Duration Scenarios

Peak	2$/w	2$/w	2$/w	2$/w	2$/w
5mins	6.9	18.88	38.76	58.64	78.53
10mins	1.64	5.6	12.2	18.8	25.4
30mins	0.32	2.3	5.6	8.9	12.2
1 hour	-0.34	0.65	2.3	3.95	5.6
4 hour	-0.83	-0.59	-0.17	0.24	0.65

mismatching handling solution for small peaks and when to use the power mismatching handling solution for large peaks. If there exists some considerable errors in the prediction, it will cause that the large power mismatching solution is misused to handle small power mismatching event or the small power mismatching solution is misused to handle the large power mismatching event, which may cause lower energy efficiency.

8.6 ESD Cost Analysis and Comparison

Cost Breakdown. Table 3 shows the cost break-down of our HHEB prototype. The energy storage devices are the most expensive components (account for 55 percent of the overall expenditure). With our existing setup, a HHEB node powers six servers and its total cost is less than 16 percent of the server total cost (approximate $4,850).

Return-On-Investment (ROI). We further simulate in light of under-provisioned power infrastructure, whether it is worth to invest hybrid energy storage to reduce CAP-EX. Similar to [6], we define the cost of procuring hybrid energy buffers to sustain e hours of peaks as e^*C_{HHEB} ($/Watt), and the CAP-EX cost of the power infrastructure to under-provision by $Ccap$ ($/Watt). The ROI for hybrid energy buffer can be calculated as: ($Ccap$ - e^*C_{HHEB}) / (e^*C_{HHEB}), where the C_{HHEB} is the total cost of SCs and batteries. We assume the battery cost C_{bat} is 300$/KWh and SC cost Csc is 10K$/KWh, as reported in [32], [37], [38]. The hybrid energy cost is: $C_{HHEB} = C_{bat}{}^*x + Csc^*y$, where x and y are the ratios of batteries and SCs and we set $x = 0.3$ and $y = 0.7$ based on our prototype. The $Ccap$ is reported to grow by $10-25 for every provisioned Watt.

We vary a wide range of $Ccap$ from 2 to 20 ($/Watt) and calculate the ROI in different peak durations as shown in Table 4. Note that the corresponding cost is amortized during the lifetime (e.g., battery: 4 years, SC: 12 years and infrastructure: 12 years). We observe a positive ROI across most of the operating regions. This suggests that deploying hybrid energy buffer is worthwhile.

Gain from Peak Shaving. Utilities often charge data-centers expensive peak cost [8]. Energy storage buffer can be used to shave peak power and save the OP-EX cost [6], [8], [32]. We assume a 100KW datacenters deployed with 20KWh homogenous batteries or hybrid energy buffer (SCs account for 30 percent and batteries account for 70 percent). The peak tariff is 12$/kW. Applying different peak shaving policies to the two types of energy buffers, we compare

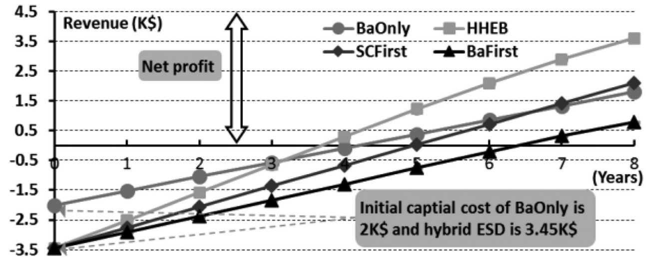


Fig. 21. The revenue comparison of different ESD power management policies (in an 8-years operation period, HHEB achieves 1.9X revenue from peak shaving benefit compared to BaOnly policy).

their revenues due to peak cost reduction within 8 years, as shown in Fig. 21. The break-even point (in year) for *BaOnly* (battery cost is 300$/KWh) is 4.2 year, similar to [8]. Taking *BaOnly* as baseline, we calculate the peak shaving gain of other three heterogeneous schemes. Our *HHEB* scheme can improve energy efficiency and reduce server downtime by 39.7 and 41 percent respectively, which are proportional to the harvested peak shaving benefit. The breakeven points of *BaFirst*, *SCFirst* and *HHEB* are 6.3, 4.9 and 3.7 years respectively. Even through the hybrid energy buffer has expensive initial CAP-EX cost than battery only buffer, with the highly efficient peak shaving policy of *HHEB*, we can earn more than 1.9X revenue from peak shaving benefit by accumulating and then averaging the per-year net profit within 8 years. On the contrary, if not appropriately managed, leveraging hybrid energy buffer may be less profitable than utilizing homogenous buffer (e.g., the net profit of *BaFirst* is less than that of *BaOnly*).

9 Conclusion

The state-of-the-art studies have proposed to shave peak power with UPS batteries for datacenters. To improve the efficiency and availability issues of previous studies, we propose a hybrid and hierarchical energy buffer scheme for datacenter to flexibly integrate various ESDs and dynamically dispatch ESD energy for handling power mismatching. We first investigate the characterizations of various ESDs. Then, we further propose HHEB, a novel energy buffering provisioning architecture that enables datacenters to deploy different ESDs in multi-hierarchies of datacenter. To efficiently utilize different energy buffers, we tailored a power management framework to intelligently and dynamically assign different ratio ESD energy to server loads for achieving higher energy efficiency and power dependability when handling power mismatching events. We further implement a scale-down prototype from scratch. We evaluate different power management policies with the prototype and the results show that HHEB could improve energy efficiency by 39.7 percent, extend UPS lifetime by 4.7X, promote energy availability by 3.2X, reduce system downtime by 41 percent and effectively improve the energy availability of various energy buffers in different hierarchies. HHEB manifests high CAP-EX ROI and is able to gain more than 1.9X peak shaving benefit during an 8-years operation period.

Acknowledgments

This work is supported in part by the National Key Research and Development Progrom of China: 2016YFB0200202, by NSF grants 1423090, 1320100, 1117261, by NSFC grant

61602368, by the Fundamental Research Funds for the Central Universities, by China Postdoctoral Science Foundation 168236, and by the Program of China Introducing Talents of Discipline to University grant B13043.

REFERENCES

[1] M. Mills, "The cloud begins with coal: big data, big networks, big infrastructure, and big power," Technical Report, Digital Power Group, *National Mining Association & American Coalition for Clean Coal Electricity*, pp. 1–45, Aug. 2013.

[2] L. Liu, H. Sun, C. Li, Y. Hu, N. Zheng, and T. Li, "Towards an adaptive multi-power-source datacenter," in *Proc. Int. Conf. Supercomputing*, Jun. 2016, Art. no. 11.

[3] U. Hoelzle and L. A. Barroso, *The Datacenters as a Computer: An Introduction to the Design of Warehouse-Scale Machines*, San Rafael, CA, USA: Morgan and Claypool Publishers, 2009.

[4] F. Xing, X. Wang, and C. Lefurgy, "How much power oversubscription is safe and allowed in datacenters," in *Proc. Int. Conf. Autonomic Comput.*, 2011, pp. 21–30.

[5] S. Pelley, D. Meisner, P. Zandevakili, T. F. Wenisch, and J. Underwood, "Power routing: Dynamic power provisioning in the datacenter," in *Proc. Int. Conf. Architectural Support Programm. Languages Operating Syst.*, 2010, pp. 231–242.

[6] S. Govindan, D. Wang, A. Sivasubramaniam, and B. Urgaonkar, "Leveraging stored energy for handling power emergencies in aggressively provisioned datacenters," in *Proc. Int. Conf. Architectural Support Programm. Languages Operating Syst.*, 2012, pp. 75–86.

[7] L. Liu, C. Li, H. Sun, Y. Hu, J. Xin, N. Zheng, and T. Li, "Leveraging heterogeneous power for improving datacenter efficiency and resiliency," *IEEE Comput. Archit. Lett.*, vol. 14, no. 1, pp. 41–45, Jan.-Jun. 2015.

[8] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, "Benefits and limitations of tapping into stored energy for datacenters," in *Proc. Int. Symp. Comput. Archit.*, 2011, pp. 341–352.

[9] N. Sharma, S. Barker, D. Irwin, and P. Shenoy, "Blink: Managing server clusters on intermittent power," in *Proc. Int. Conf. Archit. Support Programm. Languages Operating Syst.*, 2011, pp. 185–198.

[10] C. Li, Y. Hu, R. Zhou, M. Liu, L. Liu, J. Yuan, and T. Li, "Oasis: Enabling datacenter to scale out economically and sustainably," in *Proc. Int. Symp. Microarchitecture*, Dec. 2011, pp. 1960–1973.

[11] I. Goiri, R. Beauchea, K. Le, T. D. Nguyen, M. Haque, J. Guitart, J. Torres, and R. Bianchini, "GreenSlot: Scheduling energy consumption in green datacenters," in *Proc. Int. Conf. High Perform. Comput. Netw. Storage Anal.*, 2011, Art. no. 20.

[12] Z. Liu, M. Lin, A. Wierman, S. Low, and L. Andrew, "Greening geographical load balancing," *IEEE/ACM Trans. Netw.*, vol. 23, no. 2, pp. 657–671, Apr. 2015.

[13] N. Chen, X. Ren, S. Ren, and A. Wierman, "Greening multi-tenant data center demand response," in *Proc. Int. Symp. Comput. Perform. Modeling Measurements Eval.*, 2015, pp. 36–38.

[14] I. Goiri, K. Le, T. Nguyen, J. Guitart, J. Torre, and R. Bianchini, "GreenHadoop: Leveraging green energy in data processing frameworks," in *Proc. Eur. Conf. Comput. Syst.*, 2012, pp. 57–70 .

[15] W. Deng, F. Liu, H. Jin, B. Li, and D. Li, "Harnessing renewable energy in cloud datacenters: Opportunities and challenges," *Proc. IEEE Netw. Mag.*, vol. 28, no. 1, pp. 48–55, Jan. 2014.

[16] I. Goiri, W. Katsak, K. Le, T. Nguyen, and R. Bianchini, "Parasol and GreenSwitch: Managing datacenters powered by renewable energy," in *Proc. Int. Conf. Architectural Support Programm. Languages Operating Syst.*, 2013, pp. 51–64.

[17] M. Haque, K. Le, I. Goiri, R. Bianchini and T. Nguyen, "Providing green SLAs in high performance computing clouds," in *Proc. Int. Green Comput. Conf.*, 2013, pp. 1–11.

[18] L. Liu, H. Sun, C. Li, Y. Hu, T. Li, and N. Zheng. "Exploring customizable heterogeneous power distribution and management for datacenter," *IEEE Trans. Parallel Distrib. Syst.*, vol. 29, no. 12, pp. 2798–2813, Dec. 2018.

[19] [Online]. Available: http://www.microsoft.com/environment/, 2015.

[20] [Online]. Available: http://www.ibm.com/ibm/environment/climate/, 2013.

[21] [Online]. Available: http://www.apple.com/environment/renewable-energy/, 2015.

[22] [Online]. Available: http://www.google.com/green/bigpicture/, 2014.

[23] [Online]. Available: https://www8.hp.com/us/en/hp-information/social-innovation/hp-net-zero-energy-data.html/, 2013.

[24] D. Meisner, B. Gold, and T. Wenisch, "The powernap server architecture," *ACM Trans. Comput. Syst.*, vol. 29, no. 1, Feb. 2011, Art. no. 3.

[25] R. Raghavendra, P. Ranganathan, V. Talwar, Z. Wang, and X. Zhu, "No power struggles: Coordinated multi-level power management for the datacenters," in *Proc. Int. Conf. Architectural Support Programm. Languages Operating Syst.*, 2008, pp. 48–59.

[26] C. Isci, A. Buyuktosunoglu, C.-Y. Cher, P. Bose, and M. Martonosi, "An analysis of efficient multi-core global power management policies: Maximizing performance for a given power budget," in *Proc. Int. Symp. Microarchitecture*, 2006, pp. 347–358.

[27] D. Lo, L. Cheng, R, Govindaraju, L. Barroso, and C. Kozyrakis, "Towards energy proportionality for large-scale latency-critical workloads," in *Proc. Int. Symp. Comput. Architecture*, 2014, pp. 301–312.

[28] L. Liu, H. Sun, C. Li, Y. Hu, J. Xin, N. Zheng, and T. Li, "RE-UPS: An adaptive distributed energy storage system for dynamically managing solar energy in green datacenters," *J. Supercomputing*, vol. 25, no. 1, pp. 295–316, Jan. 2015.

[29] V. Kontorinis, L. Zhang, B. Aksanli, J. Sampson, H. Homayoun, E. Pettis, T. Rosing, and D. Tullsen, "Managing distributed UPS energy for effective power capping in datacenters," in *Proc. Int. Symp. Comput. Archit.*, 2012, pp. 488–499.

[30] D. Wang, C. Ren and A. Sivasubramaniam, "Virtualizing power distribution in datacenters," in *Proc. Int. Symp. Comput. Archit.*, 2013, pp. 595–606.

[31] W. Zheng, K. Ma, and X. Wang, "TE-Shave: Reducing data center capital and operating expenses with thermal energy storage," *IEEE Trans. Comput.*, vol. 64, no. 11, pp. 3278–3292, Nov. 2015.

[32] D. Wang, C. Ren, A. Sivasubramaniam, B. Urgaonkar, and H. Fathy, "Energy storage in datacenters: What, where, and how much," *SIGMETRICS Perform. Eval. Rev.*, vol. 40, no. 1, pp. 187–198, 2012.

[33] L. Narayanan, D. Wang, A. Mamun, A. Sivasubramaniam, and H. Fathy, "Should we dual-purpose energy storage in datacenters for power backup and demand response?" in *Proc. Workshop Power-Aware Comput. Syst.*, 2014, pp. 1–5.

[34] S. Schoenung and W. Hassenzahl, "Long-term vs. short-term energy storage technologies analysis: A life-cycle cost study," Rep. no.: 918358, Sandia National Laboratories, Livermore, CA, USA, 2003.

[35] M. Pedram, N. Chang, Y. Kim, and Y. Wang, "Hybrid electrical energy storage systems," in *Proc. Int. Symp. Low Power Electron. Des.*, 2010, pp. 363–368.

[36] D. Baert and A. Vervaet, "Lead-acid battery model for the derivation of peukert's law," *Electrochimica Acta*, vol. 4, pp. 3491–3504, 1999.

[37] H. Chen, T. Cong, W. Yang, C. Tan, Y. Li, and Y. Ding, "Progress in electrical energy storage system: A critical review," *Natural Sci.*, vol. 19, pp. 291–312, 2009.

[38] D. Rastler, "Electricity energy storage technology options: A white paper primer on applications, costs, and benefits," Rep. no.: 1020676, Elect. Power Res. Inst., Palo Alto, CA, USA, 2010.

[39] D. Cheng, J. Rao, C. Jiang, and X. Zhou, "Elastic power-aware resource provisioning of heterogeneous workloads in self-sustainable datacenters," *IEEE Trans. Comput.*, vol. 65, no. 2, pp. 508–521, Feb. 2016.

[40] L. Liu, H. Sun, C. Li, T. Li, J. Xin, and N. Zheng. "Managing battery aging for high energy availability in green datacenters," *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 12, pp. 3521–3536, Dec. 2017.

[41] X. Zhang and S. Reda, "Power budgeting techniques for data centers," *IEEE Trans. Comput.*, vol. 64, no. 8, pp. 2267–2278, Aug. 2015.

[42] P. Ruiu, C. Fiandrino, P. Giaccone, A. Bianco, and D. Kliazovich, "On the energy-proportionality of data center networks," *IEEE Trans. Sustainable Comput.*, vol. 2, no. 2, pp. 197–210, Apr.-Jun. 2017.

[43] M. Xu, A. Dastjerdi, and R. Buyya. "Energy efficient scheduling of cloud application components with brownout," *IEEE Trans. Sustainable Comput.*, vol. 1, no. 2, pp. 40–53, Jul.-Dec. 2016.

[44] U. Fiore, A. Castiglione, A. De Santis, and F. Palmieri, "Exploiting battery-drain vulnerabilities in mobile smart devices," *IEEE Trans. Sustainable Comput.*, vol. 2, no. 2, pp. 90–99, Apr.-Jun 2017.

[45] L. Yu, T. Jiang, Y, Cao, and Q. Qi, "Joint workload and battery scheduling with heterogeneous service delay guarantees for data center energy cost minimization," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 7, pp. 1937–1947, Jul. 2015.

[46] Y. Guo and Y. Fang, "Electricity cost saving strategy in data centers by using energy storage," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 6, pp. 1149–1160, Jun. 2013.
[47] Y. Guo, Y. Gong, Y. Fang, P. Khargonekar, and X. Geng, "Energy and network aware workload management for sustainable data centers with thermal storage," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 8, pp. 2030–2042, Aug. 2014.
[48] Y. Li, D. Wang, S. Ghose, J. Liu, S. Govindan, S. James, E. Peterson, J. Siegler, R. Ausavarungnirun, and O. Mutlu, "SizeCap: Coordinating energy storage sizing and power capping for fuel cell powered data centers," in *Proc. Int. Symp. High-Perform. Comput. Archit.*, 2016, pp. 444–456.
[49] A. Mirhoseini and F. Koushanfar, "HypoEnergy: Hybrid supercapacitor-battery power-supply optimization for energy efficiency," in *Proc. Design Autom. Test Eur. Conf. Exhib.*, 2011, pp. 1–4.
[50] L. Liu, C. Li, H. Sun, Y. Hu, J. Gu, and T. Li, "BAAT: Towards dynamically managing battery aging in green datacenters," in *Proc. IEEE/IFIP Int. Conf. Depend. Syst. Netw.*, 2015, pp. 307–318.
[51] C. Li, Y. Hu, L. Liu, J. Gu, M. Song, J. Yuan, X. Liang, and T. Li. "Towards sustainable in-situ server systems in the big data era," in *Proc. Int. Symp. Comput. Archit.*, Jun. 2015, pp. 14–26.
[52] L. Liu, C. Li, H. Sun, Y. Hu, J. Gu, T. Li, J. Xin, and N. Zheng, "HEB: Deploying and managing hybrid energy buffers for improving datacenter efficiency and economy," in *Proc. Int. Symp. Comput. Archit.*, Jun. 2015, pp. 463–475.
[53] C. Li, Z. Wang, X. Hou, H. Chen, X. Liang, and M. Guo, "Power attack defense: Securing battery-backed data centers," in *Proc. Int. Symp. Comput. Archit.*, 2016, pp. 493–505.
[54] [Online]. Available: http://www.maxwell.com/products/ultracapacitors/products/16v-large-modules, 2014.
[55] D. Linden and T. Reddy, *Handbook of Batteries*, 3rd ed. New York, NY, USA: McGraw-Hill, 2002.
[56] [Online]. Available: http://opencompute.org/, 2013.
[57] [Online]. Available: http://www.cnet.com/news/google-uncloaks-once-secret-server-10209580/, 2014.
[58] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, Hoboken, NJ, USA: Wiley, 2013.
[59] P. Goodwin, "The holt-winters approach to exponential smoothing: 50 years old and going strong," *Int. J. Appl. Forecasting*, vol. 59, pp. 30–33, 2010.
[60] S. Huang, J. Huang, J. Dai, T. Xie, and B. Huang, "The HiBench benchmark suite: Characterization of the MapReduce-based data analysis," in *Proc. Data Eng. Workshops Int. Conf. Data Eng.*, 2010, pp. 209–228.
[61] The CloudSuite 2.0, [Online]. Available: http://parsa.epfl.ch/cloudsuite, 2014.
[62] H. Bindner, T. Cronin, P. Lundsager, J. Manwell, U. Abdulwahid, and I. Gould, "Lifetime modelling of lead acid batteries," Risø National Laboratory, Roskilde, Denmark, Rep. no. ENK6-CT-2001-80576, 2005.

Longjun Liu received the PhD degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2015. He was a visiting PhD student in the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, from 2012 to 2014. Currently, he is an assistant professor with the Xi'an Jiaotong University. His research interests include computer architecture, power-efficient and intelligent computing systems, and energy management for green dataenters. He is a member of the IEEE.

Hongbin Sun (M'11) received the BS and PhD degrees in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 2003 and 2009, respectively. He was a visiting PhD student in the Electrical, Computer, and Systems Engineering Department at Rensselaer Polytechnic Institute, Troy, NY, from 2007 to 2008. Currently, he is an associate professor with Xi'an Jiaotong University. His research interests include memory hierarchy design and VLSI architecture for digital video processing. He is a member of the IEEE.

Chao Li received the PhD degree from the University of Florida, in 2014. He is an assistant professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University. His research focuses on computer architecture, energy-efficient systems, and green computing. He is a member of the IEEE.

Tao Li received the PhD degree in computer engineering from the University of Texas at Austin. He is a professor with the Department of Electrical and Computer Engineering, University of Florida, USA. His current research interests include computer architecture, microprocessor, memory storage system design, virtualization technologies, energy efficient, sustainable, dependable data centers, cloud/big data computing platforms, the impacts of emerging technologies/applications on computing, and evaluation of computer systems. He is a fellow of the IEEE.

Jingmin Xin (S'92–M'96–SM'06) received the BE degree in information and control engineering from Xi'an Jiaotong University, Xi'an, China, in 1988, and the MS and PhD degrees in electrical engineering from Keio University, Yokohama, Japan, in 1993 and 1996, respectively. Since 2007, he has been a professor at Xi'an Jiaotong University. His research interests include the areas of adaptive filtering, statistical and array signal processing, system identification, and pattern recognition. He is a senior member of the IEEE.

Nanning Zheng (SM'93–F'06) received the graduate degree from the Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an, China, in 1975, the MS degree in information and control engineering from Xi'an Jiaotong University, in 1981, and the PhD degree in electrical engineering from Keio University, Yokohama, Japan, in 1985. He jointed Xi'an Jiaotong University in 1975, and is currently a professor and the director of the Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University. His research interests include computer vision, pattern recognition and image processing, and hardware implementation of intelligent systems. He is a fellow of the IEEE.