IEEE

COMPUTER ARCHITECTURE LETTERS

January-June 2015 | Volume 14 | Number 1 | ICALC3 | (ISSN 1556-6056)

A publication of the IEEE Computer Society
Originated by TCCA, November, 2001
Indexed in ISI

PAPERS

(Contents continued on back cover)

IEEE computer society

IEEE

www.computer.org

Leveraging Heterogeneous Power for Improving Datacenter Efficiency and Resiliency

Longjun Liu, Chao Li, Hongbin Sun, Yang Hu, Jingmin Xin, Nanning Zheng, and Tao Li

Abstract—Power mismatching between supply and demand has emerged as a top issue in modern datacenters that are under-provisioned or powered by intermittent power supplies. Recent proposals are primarily limited to leveraging uninterruptible power supplies (UPS) to handle power mismatching, and therefore lack the capability of efficiently handling the irregular peak power mismatches. In this paper we propose hPower, the first heterogeneous energy buffering strategy that incorporates supercapacitors into existing datacenters to handle power mismatch. Our technique exploits power supply diversity and smart load assignment to provide efficiency-aware and emergency-aware power mismatch management. We show that hPower could improve energy efficiency by 30 percent, extend UPS lifetime by 4.3×, and reduce system downtime by 36 percent. It allows datacenters to adapt themselves to various power supply anomalies, thereby improving operational efficiency and resiliency.

Index Terms—Energy-aware systems, performance of systems, computer system implementation

1 Introduction

MATCHING power supply to instantaneous demand is crucial to eliminate service disruption and maintain desired power efficiency in datacenters. However, a power mismatching issue arises as many datacenters today start to over-subscribe power infrastructure to reduce power demand charge from the utility grid [1], [2], or integrate renewable energy systems to cap IT carbon footprint [3], [4]. For example, unexpected load surge can easily generate power peaks that exceed the maximum power budget, leading to costly downtime in an under-provisioned datacenter. The unstable renewable power generation can cause varying degrees of demand-supply power mismatches, resulting in significant adverse impact on efficiency and availability.

To overcome the power mismatching issue, recent studies have started to aggressively explore the energy backup offered by conventional UPS batteries [2], [5]. However, such design incurs frequent charging/discharging cycles caused by various power supply anomalies, which can greatly compromise UPS lifetime. In addition, due to the limitations of the internal diffusion phenomenon, batteries incur very fast and significant capacity drop at peak current [6]. This poses serious threat to datacenter availability during peak hours.

We propose *hPower*, which explores the benefits of introducing hybrid energy buffering into datacenters. Specifically, we integrate supercapacitors (SC) (a.k.a ultracapacitor) with conventional UPS systems to provide an additional layer of safety in the event of an unexpected peak power mismatch. Supercapacitors have emerged as a promising alternative to electrochemical batteries despite their relatively higher price. They are known to have very low round-trip energy loss and normally two to three orders of magnitude more life cycles than batteries [7]. A hybrid energy buffering system allows us to deliver high levels of peak current very quickly without draining the main UPS systems.

When we transition from homogeneous to heterogeneous energy buffers, challenge arises as the latter require careful power assignment between batteries and SCs. For a given power demand value, there is an optimal way of discharging that could provide the longest discharging duration. This optimal discharging point often moves when the available stored energy changes in either batteries or SCs. In addition, from the viewpoint of energy efficiency, the ideal usage pattern of heterogeneous energy buffers also depends on the power mismatch value. For example, when the power peak is small and narrow, it is better to only use SC to provide power shortfall. This is because SCs can be easily charged in a short time with negligible round-trip energy loss.

Contributions: 1) We explore key design considerations of a promising power delivery scheme: using heterogeneous energy buffers to handle power mismatching. 2) We propose hPower, a novel power management strategy for maximizing the benefits of heterogeneous energy buffers in datacenters.

2 Heterogeneous Energy Buffers: Overview and Key Design Considerations

With heterogeneous energy buffers, one can overcome the limitations of relying on any single type of storage devices, thereby achieving better energy reliability and efficiency. In this study we mainly focus on a heterogeneous energy buffer that integrates supercapacitors with lead-acid batteries.

2.1 Advantages of Supercapacitors

One of the primary reasons for using SCs to buffer energy is that they incur negligible round-trip energy loss [8]. Our real measurements indicate that SCs can achieve 93 $\sim$ 95 percent round-trip energy efficiency, as shown in Fig. 1. In contrast, lead-acid batteries (widely used in UPS systems) have less than 80 percent efficiency even in the best case in our experiment.

In fact, the efficiency of batteries can be even lower 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 discharge - part of the stored energy seems to be "lost"; during periods of no or very low discharge, they can recover the energy "lost" to a certain extent [6], [10]. Fig. 1 shows our characterization of different discharging methods. The one-time discharging efficiency of the lead-acid battery decreases as we add more servers to increase the power demand. Given additional discharge cycles and enough recovery time, the battery efficiency can increase significantly, by 6 $\sim$ 24 percent. However, it 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 also be very large, which account for almost half of the recovered energy, as shown in Fig. 1.

Another reason for choosing SCs is that they have excep-tional cycle life. Although SC has high capital cost, the amortized cost of SCs to each charging and discharging cycle (Capital cost/Lifetime cycle, \$/KWh/cycle) is competitive to batteries. In Fig. 2, the initial cost of most UPS batteries is about 100-150 \$/KWh, while for SCs it is about 16K $\sim$ 30K \$/KWh [7-9]. However, the amortized cost of SCs is close to NiCd and Li-ion batteries (about 0.4 \$/KWh per cycle).

2.2 Power Management Challenges

By using a supercapacitor, designers can deliver the high current levels needed for these power mismatching events and then recharge quickly between events. However, it does not necessarily mean that one should always give SC systems high priority when shaving power peaks.

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

Manuscript received 10 Aug. 2014; revised 23 Sept. 2014; accepted 7 Oct. 2014. Date of publication 14 Oct. 2014; date of current version 19 June 2015.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/LCA.2014.2363084

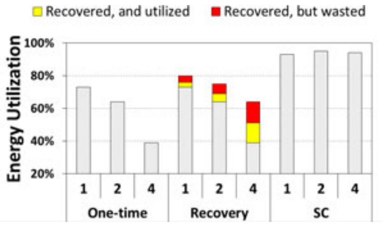

Fig. 1. Energy efficiency measurement using different numbers of servers.

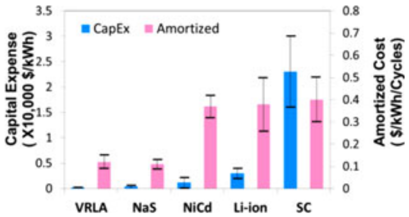

Fig. 2. Cost comparison of different energy storage technologies.

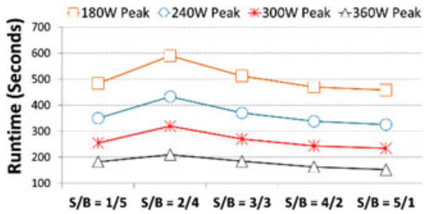

Fig. 3. Discharge duration (*m*/*n* means *m* servers for SC and *n* servers for battery).

We experiment with our heterogeneous energy buffering system (190Wh Battery, 20Wh SC) and a cluster of six servers running synthetic workloads. We vary the number of servers assigned to the battery system and SC system to test the maximum server runtime under constant power demand. In the experiment, whenever one energy storage device is depleted, the other will take over the entire load immediately using power switches. As Fig. 3 shows, there is an optimal load assignment that can provide the longest discharging time. It is clear that one should not heavily rely on either supercapacitors or batteries. For example, by assigning heavy load on supercapacitors, the server cluster runtime (uptime) can be decreased by 25 percent on average.

Note that there is not a fixed optimal operating point. The optimal server assignment actually depends on the capacity of energy buffers and the shape of power peaks. For each power mismatch event, the system must first evaluate whether or not the maximal discharge duration is important. For these short duration peaks, it is preferable to only use high-efficiency SCs to supply power. If the power peak is wide and high, one must evaluate the available energy stored in SCs and batteries and determine the desired power allocation scheme to minimize server downtime.

3 HPOWER POWER DESIGN SCHEME

SCs are well complementary to a battery. They typically have limited energy capacity but are able to handle power surges. This is why they are used in conjunction with each other. In this section we introduce hPower, our design of heterogeneous energy buffering system in data centers.

3.1 System Architecture

Centralized battery and distributed battery are two exiting energy storage architecture in datacenters. Fig. 4 compares our heterogeneous system to prior work. Conventional centralized battery system is not scalable and it only provides load shifting for the entire datacenter. The emerging distributed battery design is scalable and

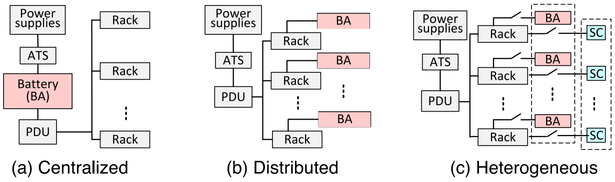

Fig. 4. A comparison of different energy buffer architecture.

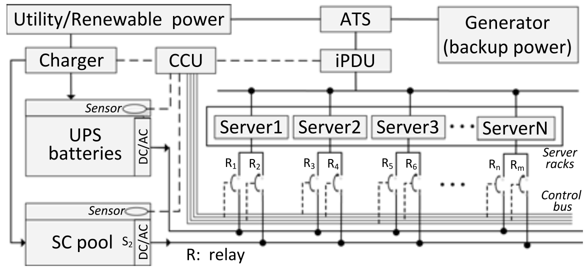

Fig. 5. Overview of the hPower power architecture design.

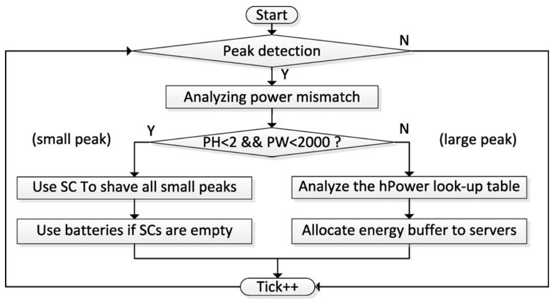

Fig. 6. The flowchart of hPower dynamic load assignment.

allows datacenters to shave power peaks using a fraction of the installed battery.

Our heterogeneous energy buffer design extends prior work in two aspects. First, in addition to battery cabinets, each server rack can access to a SC module and its power delivery path is controlled via distributed power switches. In Fig. 5, the power switch based battery control is enlightened by the reconfigurable battery array in the power system community. It enables a datacenter to dynamically determine the distribution of server power demand between batteries and supercapacitors. Second, batteries are further interconnected to a battery pool and so do 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 peak power mismatch since they can be charged and discharged very quickly.

3.2 Dynamic Load Assignment

During runtime, hPower dynamically distributed load power demand between batteries and SCs. Fig. 6 shows the flowchart for managing power mismatches.

Our system treats batteries and SCs as a unified energy buffer when the power mismatch is significant and the dura-tion is long. In other words, we use batteries and SCs to jointly shave large power peaks. To determine the desired operating point, our system maintains a power allocation table (PAT) for its heterogeneous energy buffer. This table specifies the ideal power assignment between batteries and SCs. Each entry of the power allocation table is indexed by the available energy levels of the battery and the SC pool. The initial value of this table is obtained via profiling in a pilot scheme like Fig. 3. We dynamically control the on/off of the distributed power switches based on the specified server assignment ratio in the PAT. This allows use to achieve the ideal discharging rate and the longest discharging duration.

```
Obtain current SC capacity: SC_initial , Battery capacity: BA_initial ;
For table i = 1 to n   // search the look-up table PAT
    If (SC_index == SC_initial && BA_index == BA_initial)
        index = i;
If (index == 0)  //does not find a matched entry
    index = SeekSimilar(SC_initial , BA_initial); //search the most similar capacity pair
Server ratio R = PAT_index;
Allocate different numbers of servers to SC and BA based on R;
Collect running results at the end of the power mismatch event.
If (SC_end/BA_end > SC_initial/BA_initial)
    R = R + Δr; //SC receives increased server assignment
Else If (SC_end/BA_end < SC_initial/BA_initial)
    R = R - Δr; //BA receives increased server assignment
Update {SC_initial, BA_initial, ΔR} in the PAT look-up table
```

Fig. 7. Pseudo code for managing the power allocation.

TABLE 1
Evaluated Power Schemes

Schemes	Architecture
BaOnly	Battery only
BaFirst	Hybrid (battery first use)
ScFirst	Hybrid (SC first use)
hPower	Hybrid (dynamically use)

TABLE 2
Key Parameters Used in Simulation

Parameters	Values (SCs/Batteries) [5]–[9]
Capacity	12Ah per server/28Ah per server
Efficiency	95%/83%
Life cycle	500,000 cycles/5,000 cycles
Initial Cost	20K$/KWh/130$/KWh

When the height of power surge is mild and the duration is short, hPower threat battery and supercapacitors as a two-tier energy storage system. It will use SCs to supply small and frequent power mismatch events and charge SCs between the events. This is because supercapacitors have much better round-trip energy efficiency and they can be charged and discharged continuously without degradation.

Whatever large or small peaks, SCs are charged firstly between these peak mismatch events for the high energy efficiency. After all the SCs are charged to their full capacity, batteries are charged alternatively during the following times.

3.3 Optimizing Energy Buffering

The PAT table does not always guarantee the optimal discharging results. This is because the initial profiling data stored in it based on a pilot run can be less accurate. On the other hand, this reflects the uncertainty surrounding possible battery wear-out and any kind of capacity degradation that may occur in the near future.

To ensure continuous control effectiveness, hPower can self-optimize the load assignment over its lifetime. As shown in Fig. 7, the load ratio R determines how much percentage of server power demand is allocated to supercapacitors. For every discharging event it manages, our system monitors the remaining capacity in SCs and batteries after discharging. If the actual battery capacity decline rate is faster than expected (e.g., due to internal wear-out and self-discharging), hPower will increase the load ratio by $\Delta r = 1\%$ to increase the usage of SCs in the next control period. Otherwise, hPower will decrease the load ratio. The algorithem needs to keep track of the energy buffer's effective capacity so that the R is mapped to a closely matched assignment. Our solution is orthogonal to system-level power management schemes and is complementary to existing power coordination framework like [14].

TABLE 3
Technical Data of The Evaluated Datacenter Traces

Workload	Trace [12], [13]	Peak height	Peak width	Peak freq.
HWD	LLNL-T3D-1996-2	2.51 (High)	2,384 (Wide)	141 (Dense)
HWS	HPC2N-2002-2.1	2.16 (High)	4,279 (Wide)	33 (Sparse)
HND	NASA-iPSC-1993-3.1	2.62 (High)	1,418 (Narrow)	246 (Dense)
HNS	NASA-HTTP	2.12 (High)	265 (Narrow)	80 (Sparse)
LWD	SDSC-BLUE-2000-4.1	1.64 (Low)	4,021 (Wide)	115 (Dense)
LWS	CTC-SP2-1996-3.1	1.43 (Low)	5,930 (Wide)	65 (Sparse)
LND	LANL-O2K-1999-2	1.32 (Low)	1,886 (Narrow)	218 (Dense)
LNS	LBL-CONN-7	1.52 (Low)	221 (Narrow)	56 (Sparse)

(The threshold that separates high peak and low peak is $PH = 2$; the threshold that separates wide peak and narrow peak is $PW = 2{,}000$ seconds; the threshold that separates dense peaks and sparse peaks is $PF > 100$ times per week).

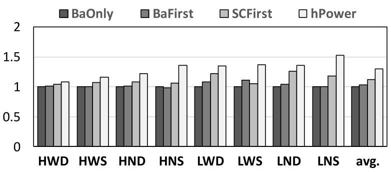

Fig. 8. Normalized energy efficiency under different power management schemes.

4 Evaluation Methodology

We develop a discrete-event simulation framework for datacenters that use heterogeneous energy buffers. It can process a chronological sequence of job submissions and accurately calculate the power mismatch value. We maintain detailed running logs of each event to determine the remaining capacity and cycle life of energy buffers. Detailed evaluation framework and methodology are discussed in [11].

Table 1 summarizes the evaluated peak power management schemes. BaOnly is a conventional scheme that only uses UPS batteries to shave peak power. BaFirst and ScFirst both use hybrid energy storage devices, i.e., a combination of SC and batteries. However, they do not have intelligent server allocation and use a priority based discharging method. Table 2 shows the key parameters we used in our simulation.

We evaluate real-world datacenter traces collected from well-established online repositories [12], [13]. They include both HPC datacenter traces and internet server traces. As shown in Table 3, we classify different traces based on the height (PH), width (PW), and frequency (PF) of power peaks.

5 Results

To improve the energy efficiency of heterogeneous energy buffers, one must fully utilize both super-capacitors and batteries. In Fig. 8 we first show the overall energy efficiency measurement. Compared to a conventional battery-only power scheme, the utilization improvement for BaFirst, ScFirst, and hPower is 3, 12, and 30 percent, respectively. The reason why BaFirst is very close to a battey-only design is that BaFirst rarely use its super-capacitors. If we always discharge the SC first, we can greatly reduce energy loss. However, when the SCs are depleted, batteries will have to handle all the high current drawn. This can cause significant efficiency degradation. In contrast, hPower can balance the usage of the two, and therefore has the highest efficiency. Another key benefit of hPower is that it improves server availability. In Fig. 9 we evaluate the server downtime of different energy buffering schemes. The server downtime here is the aggregated time when the server power demand exceeds the power budget but the energy buffers do not have enough power to shave the power peak. The normalized server downtime for BaFirst, ScFirst, and hPower is

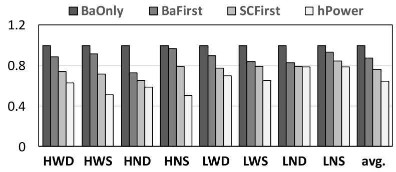

Fig. 9. Normalized server downtime under different power management schemes.

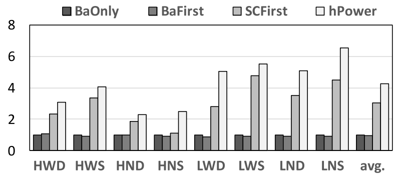

Fig. 10. Normalized Battery lifetime of different power management schemes.

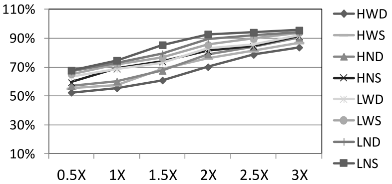

Fig. 11. Energy efficiency under different total energy buffer capacity.

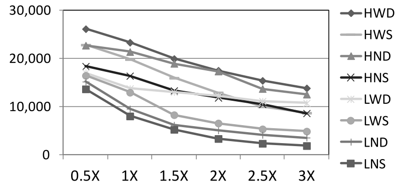

Fig. 12. Server downtime under different total energy buffer capacity.

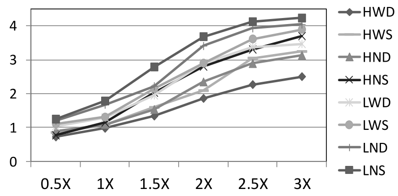

Fig. 13. Battery lifetime comparison under different total energy buffer capacity.

87, 76, and 64 percent, respectively. During runtime, hPower can always maintain the longest discharging duration by dynamically adjusting the server assignment between SCs and batteries. This is especially useful for handling high and wide power peaks.

We use the Ah-Throughput battery lifetime model [15] to caculate the anticipated battery lifetime based on detailed battery usage logs. As shown in Fig. 10. It shows that hPower can improve battery life by 4.3×, while ScFirst only shows 3× lifetime improvement. In a heterogeneous energy buffering system, the lifetime of batteries is the bottleneck of system lifespan and requires special attention. Carefully managing the usage of the battery and SC system allows us to maximize the potential of heterogeneous energy buffers.

In Figs. 11, 12, and 13, we change the total installed capacity of energy buffers and measure the absolute value of efficiency, server downtime, and battery cycle life. It's clear that larger capacity can improve the efficiency and system resilience. But the relationship between performance and capacity may not be linear. To achieve the best cost-effectiveness, a more thorough capacity planning analysis is necessary.

7 Related Work

Recent efforts have focused on repurposing UPS batteries [2], [5] to address peak power in datacenters. Energy storage is also the key enabler for integrating alternative energy sources into the energy portfolio of datacenters [3], [4], [16]. These works can effectively shave power spikes but only use batteries as the sole tuning knob. Energy efficiency and battery life can be a major problem for these designs. While SCs have grabbed certain attention in prior characterization work [17], [18], their power architecture and management schemes for power mismatching shaving have not been studied in depth to date. Distinguished from prior work, this paper makes the first step towards designing a more efficient and resilient energy buffering system.

8 Conclusions

In this study, batteries and SCs are pooled and delivered in datacenters as heterogeneous energy buffers. We show that heterogeneous power scheme is about more than a hybrid energy backup system. It demands very careful energy capacity monitoring and smart power demand allocation. Datacenters supported by such a heterogeneous energy buffering system are shown to have much better energy efficiency and peak power handling capability.

Acknowledgment

This work was supported in part by NSF grants 1423090, 1117261, by the National Basic Research Program of China (973 Program) 2010CB327902, and by NSFC grant 61274028. Longjun Liu was also supported by the China Scholarship Council for one year study at the University of Florida. Hongbin Sun is the corresponding author.

References

[1] X. Fan, W.-D. Weber, and L. A. Barroso, "Power provisioning for a warehouse-sized computer," in *Proc. 36th Annu. Int. Symp. Comput. Archit.*, 2007, pp. 13–23.
[2] S. Govidan, D. Wang, A. Sivasubramaniam, and B. Urgaonkar, "Leveraging stored energy for handling power emergencies in aggressively provisioned datacenters," in *Proc. 17th Int. Conf. Archit. Support Programm. Lang. Oper. Syst.*, 2012, pp. 75–86.
[3] C. Li, Y. Hu, R. Zhou, M. Liu, L. Liu, J. Yuan, and T. Li, "Oasis: Enabling datacenter servers to scale out economically and sustainably," in *Proc. 46th Annu. Int. Symp. Microarchit.*, 2013, pp. 322–333.
[4] I. Goiri, W. Katsak, K. Le, T. D. Nguyen, and R. Bianchini, "Parasol and greenswitch: Managing datacenters powered by renewable energy," in *Proc. 17th Int. Conf. Archit. Support Programm. Lang. Oper. Syst.*, 2013, pp. 51–64.
[5] V. Kontorinis, L. Zhang, B. Aksanli, J. Sampson, H. Homayoun, E. Pettis, D. M. Tullsen, and T. S. Rosing, "Managing distributed ups energy for effective power capping in data centers," in *Proc. 36th Annu. Int. Symp. Comput. Archit.*, 2012, pp. 488–499.
[6] D. Linden, *Handbook of Batteries*, 3rd ed. New York, NY, USA: McGraw-Hill, 2002
[7] S. Schoenung and W. V. Hassenzahl, "Long-vs. short-term energy storage technologies analysis," Sandia Nat. Laboratories, Albuquerque, NM, USA, Tech. Rep. SAND2003-2783, 2003.
[8] H. Chen, T. N. Cong, W. Yang, C. Tan, Y. Li, and Y. Ding, "Progress in electrical energy storage system: A critical review," *Progress Natural Sci.*, vol. 19, no. 3, pp. 291–312, 2009.
[9] D. Rastler, "Electricity energy storage technology options: A white paper primer on applications, costs, and benefits," Electric Power Res. Inst., Palo Alto, CA, USA, Tech. Rep. 000000000001020676, 2010.
[10] H. Kim and K. Shin, "Scheduling of battery charge, discharge, and rest," in *Proc. Real-Time Syst. Symp.*, 2009, pp. 13–22.
[11] C. Li, R. Wang, T. Li, D. Qian, and J. Yuan, "Managing green datacenters powered by hybrid renewable energy systems," in *Proc. 11th Int. Conf. Auton. Comput.*, 2014, pp. 261–272.
[12] [Online]. Available: http://www.cs.huji.ac.il/labs/parallel/workload/logs.html
[13] The Internet Traffic Archive, [Online]. Available: http://ita.ee.lbl.gov/
[14] R. Raghavendra, P. Ranganathan, V. Talwar, Z. Wang, and X. Zhu, "No power struggles: Coordinated multi-level power management for the data center," in *Proc. 17th Int. Conf. Archit. Support Programm. Lang. Oper. Syst*, 2008, pp. 48–59.
[15] H. Bindner, T. Cronin, P. Lundsager, J. F. Manwell, U. Abdulwahid, and I. Baring-Gould, "Lifetime modelling of lead acid batteries," Risø Nat. Laboratory, Roskilde, Denmark, Tech. Rep., 2005.

[16] C. Li, R. Zhou, and T. Li, "Enabling distributed generation powered sustainable high-performance data center," in *Proc. 19th Int. Symp. High Perform. Comput. Archit.*, 2013, pp. 35–46.

[17] D. Wang, C. Ren, A. Sivasubramaniam, B. Urgaonkar, and H. Fathy, "Energy storage in datacenters: What, where, and how much?" in *Proc. 12th ACM Sigmetrics/Perform. Joint Int. Conf. Meas. Model. Comput. Syst.*, 2012, pp. 187–198.

[18] M. Pedram, N. Chang, Y. Kim, and Y. Wang, "Hybrid electrical energy storage systems," in *Proc. 16th ACM/IEEE Int. Symp. Low Power Electron. Design*, 2010, pp. 363–368.