Managing Green Datacenters Powered by Hybrid Renewable Energy Systems

Chao Li[1], Rui Wang[2], Tao Li[1], Depei Qian[2], and Jingling Yuan[3]
[1]*University of Florida* [2]*Beihang University* [3]*Wuhan University of Technology*

Abstract

The rapidly growing server energy expenditure and the warning of climate change have forced the IT industry to look at datacenters powered by renewable energy. Existing proposals on this issue yield sub-optimal performance as they typically assume certain specific type of renewable energy sources and overlook the benefits of cross-source coordination. This paper takes the first step toward exploring green datacenters powered by hybrid renewable energy systems that include baseload power supply, intermittent power supply, and backup energy storage. We propose *GreenWorks*, a power management framework for green high-performance computing datacenters powered by renewable energy mix. Specifically, GreenWorks features a hierarchical power coordination scheme tailored to the timing and capacity of different renewable energy sources. Using real datacenter workload traces and renewable power generation data, we show that our scheme allows green datacenters to achieve better design trade-offs.

1. Introduction

The global server power demand reaches approximately 30 gigawatts in total [1], which account for over 250 million metric tons of CO_2 emissions per year [2]. Faced with a growing concern about the projected rise in both server power demand and carbon emissions, academia and industry alike are now focusing more attention than ever on non-conventional power provisioning solutions. For instance, recently there have been vigorous discussions on renewable energy driven computer system design with respect to carbon-aware scheduling [3-5], renewable power control [6-9], and cost optimization strategies [10-11]. In addition, Microsoft, eBay, HP, and Apple have announced projects that use green energy sources like solar/wind power, fuel cells, and bio-gas turbines to minimize their reliance on conventional utility power [12-15]. It has been estimated that these eco-friendly IT solutions could reduce almost 15% global CO_2 emissions by 2020, leading to around $900 billion of cost savings [16].

The expected growth in renewable power generation poses new challenges for datacenter operational resilience. A number of the renewable energy sources are *intermittent power supply*, such as wind turbine and solar array. They are free sources of energy but incur power variability problems. Several emerging green power supplies, such as fuel cells and bio-fuel based generators, are typically used as *baseload power supply*. They are stable and controllable power sources, but not fast enough to respond instantaneously to quick changes in server power demand. In case the intermittent power supply drops suddenly or the baseload power supply cannot follow an unexpected power demand surge, *backup power supply* (e.g., batteries, super-capacitors) must be used to handle the power shortfall.

As we move toward a smarter grid, datacenters are expected to be powered by hybrid renewable energy systems that combine multiple power generation mechanisms [17]. With an integrated mix of complementary power provisioning methods, one can overcome the limitations of each single type of power supply, thereby achieving better energy reliability and efficiency.

However, a common limitation of prior proposals is that they mainly focus on certain specific type of green power supplies. We classify existing schemes into three broad categories: 1) *load shedding*, which focuses on utilizing intermittent power [6, 9], typically reduces load when renewable power drops; 2) *load boosting*, which uses both intermittent and backup power [8, 18], takes advantage of the stored energy to maintain desired performance when the current green power generation is inadequate; and 3) *load following*, which assumes both baseload and backup power [19], leverages tunable generators to track datacenter load demand. Since prior proposals lack the capability of managing renewable energy mix, they can hardly gain the maximum benefits from hybrid renewable energy systems, and consequently yield sub-optimal design tradeoffs.

In this study we explore diversified multi-source power provisioning for green high-performance datacenters today and in the future. We propose *GreenWorks*, a framework for managing datacenter power across several layers from datacenter server to onsite renewable energy mix. GreenWorks comprises two key elements: the *green workers*, which are multiple platform-specific power optimization modules that use different supply/load control strategies for different types of renewable energy systems; and *green manager*, a hierarchical coordination scheme for green workers.

GreenWorks tackles the challenges of integrating and coordinating heterogeneous power supplies with a three-tiered hierarchical coordination scheme. Each layer of the hierarchy is tailored to the specific timing and utilization requirements of the associated energy sources. In addition, power management modules in different layers of the hierarchy can also interact with each other within the framework. This allows us to further improve the power management effectiveness of hybrid renewable energy systems.

GreenWorks emphasize a multi-objective power management. It jointly manages green energy utilization, backup energy availability, and workload performance. Specifically, we define three types of green workers: 1) *baseload laborer*, which adjusts the output of the baseload power to track the coarse-grained changes in load power demand; 2) *energy keeper*, which regulates the use of the stored renewable energy to achieve satisfactory workload performance while maintaining desired battery life; and 3) *load broker*, which could opportunistically increase the server processing speed to take advantage of the excess energy generation of the intermittent power supply. All the three modules are able to distill crucial runtime power profiling data and identify appropriate control strategies for different types of renewable generation.

To our knowledge, this paper is the first to design a hierarchal power management and coordination framework for multi-source powered green datacenters.

This paper makes three main contributions:

- We propose GreenWorks, a hierarchical power management framework for green datacenters powered by renewable energy mix. It enables cross-source power management coordination, thereby greatly facilitating supply-load power matching.
- We propose a multi-source driven multi-objective power management that takes advantage of our hierarchical power management framework. Our technique enables GreenWorks to maximize the benefits of the hybrid renewable energy systems without heavily relying on any single type of power supply.
- We evaluate GreenWorks using real-world workload traces and green energy data. We show that GreenWorks could achieve less than 3% job runtime increase, extend battery lifetime by 23%, increase UPS backup time by 12%, and maintains the same energy efficiency as the state-of-the-art design.

The rest of this paper is organized as follows. Section 2 introduces background. Section 3 proposes the GreenWorks framework. Section 4 proposes multi-objective power management scheme. Section 5 describes evaluation methodologies. Section 6 presents our results. Section 7 discusses related work and Section 8 concludes this paper.

2. Background

Today's energy crisis and environmental problems force the IT industry to look at datacenter power provisioning in a different way. In this section, we introduce green datacenters powered by hybrid renewable energy systems and discuss their design challenges.

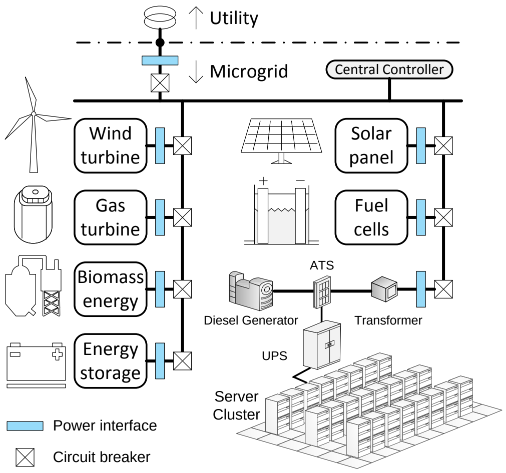


Figure 1: A datacenter powered by renewable power mix.

2.1 Hybrid Renewable Energy Systems

There are three types of renewable power supplies that we can leverage to power a datacenter. Some green power supplies, such as solar panels and wind turbines, are affected by the availability of ambient natural resources (i.e., solar irradiance or wind speed). They are referred to as *intermittent power supply* since their outputs are time-varying. Several emerging green power supplies – including fuel cells, bio-fuel based gas turbines, and bio-mass generators – can offer controllable green energy by burning various green fuels. We refer to them as *baseload power supply* since they can be used to provide stable renewable power to meet the basic datacenter power demand (e.g., idle power). In addition, energy storage devices such as batteries and super-capacitors are also critical components that provide *backup power supply*. They can be used to temporarily store green energy or improve power quality.

Looking ahead, datacenters in the smart grid era are expected to be powered by *hybrid renewable energy systems* that combine all the three types of power supplies, as shown in Figure 1. Different power supplies are typically implemented as small, modular electric generators (called micro-sources) near the point of use. To manage such an integrated renewable energy mix, micro-grid is proposed as a coordinated cluster/network of supply and load [20]. Although the micro-grid allows its customer to import power from the utility, we focus our attention on minimizing the reliance on utility power due to sustainability and cost concerns.

Energy source management and datacenter load management are largely decoupled in prior studies. Existing micro-grid control strategies often focus on power supply scheduling [21]. Recent proposals on power-aware datacenter mainly emphasize demand response control [22, 23]. In contrast, we propose load/supply cooperative power management across several layers from servers to hybrid renewable energy systems.

Micro-sources	Response Speed	Startup Time
Batteries	Immediate	N/A
Flywheel	Immediate	N/A
Fuel cells	30 sec ~ 5 min	20 ~ 50 min
Gas turbine	10s of seconds	2 ~ 10 min

Table 1: Response speed of different power supplies [24-28].

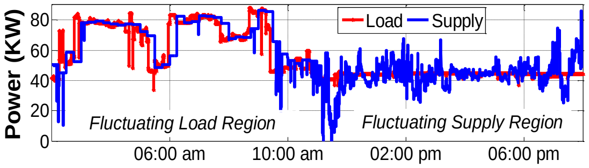


Figure 2: The demand-supply mismatch scenario.

2.3 Energy Balance Challenge

Many system-level events can cause power demand fluctuations, such as dynamic power tuning via DVFS, on/off server power cycles, and random user request. Unexpected variations in intermittent power supplies, unfavorably combined with datacenter workload fluctuation, could make the power mismatch problem even worse. Therefore, matching datacenter load to the variable power budget is often the crux of eliminating power disruptions in a green datacenter.

Managing multi-source powered system can be a great undertaking. As shown in Table 1, micro-sources often have different characteristics and operating timeframes. Most baseload green power systems cannot meet the needs of fast supply-load power matching. For example, both fuel cells and gas turbines need time to be committed and dispatched to a desired output level. They provide a slow energy balance service called *load following*, which typically occurs every tens of minutes to a few hours [28].

Figure 2 illustrates the load matching effectiveness using real-world datacenter traces and renewable energy datasets. The power supply trace shown in the figure combines the outputs of baseload power units and intermittent wind power supplies. As can be seen, load following alone cannot eliminate fine-grained power mismatch. When the wind power is stable, fluctuating load can be the main cause of power mismatch; when wind power output varies, it can significantly increase mismatch events. Although increasing the baseload power output can reduce the chance of brownout, it will significantly increase the operational expenditure.

Note that we cannot heavily rely on utility power grid and energy backup to manage the demand-supply power mismatches. First, it requires additional standby power capacity, which is economically unfavorable. Energy backup services are typically much more costly than the load following services [29]. Second, grid-inverter and battery incur round-trip energy loss, which degrade overall system efficiency. Third, heavy reliance on backup power supply can be risky. As recent survey indicates, datacenters in the US experiences 3.5 times of utility power loss per year with an average duration over 1.5 hours [30]. It also shows that UPS battery failure and capacity exceeded are the top root causes of unplanned outages. Without appropriate coordination, the demand-supply power mismatch can cause frequent battery discharging activities, which not only decrease the battery lifetime but also frequently deplete the stored energy that is crucial for handling emergencies.

In this study we explore a holistic approach for eliminating the supply-load mismatch problems in green datacenters. Specifically, we look at how cross-source power management and coordination will help to improve energy balance and datacenter resilience.

3. The GreenWorks Framework

GreenWorks is a hierarchical power management scheme that is tailored to the specific timing and utilization requirements of different energy sources. It provides coordinated power management across intermittent renewable power supplies, controllable baseload generators, onsite batteries, and datacenter servers.

The intention of this work is to provide an initial power management framework for datacenters powered by renewable energy mix. In the smart grid era, datacenters must increase their awareness of the attributes of power supplies to achieve the best design trade-offs.

3.1 System Overview

Figure 3 depicts the architecture of a green datacenter powered by renewable energy mix. We adopt typical micro-grid power distribution scheme for managing various renewable energy resources. Various renewable energy systems are connected to the power feeder through circuit breakers and appropriate interfaces.

GreenWorks is a middleware that resides between front-end computing facilities and back-end distributed generators. It manages various onsite energy sources through a micro-grid central controller, which is a typical power management module in the micro-grid system. The controller is able to adjust onsite power generation through communication with the dedicated power interface connected to each distributed generators. GreenWorks also communicates with the UPS battery rack, the cluster-lever power meters, and the server-level power control module. It cooperatively adjusts power supplies and workload performance levels, and thereby eliminates demand-supply power mismatch.

As shown in Figure 3, GreenWorks comprises two key elements: the *green workers* and the *green manager*. The former are platform-specific power management modules for managing different types of micro-sources and the later coordinates these modules. In this study we define three types of green workers: *baseload laborer* (B), *energy keeper* (E), and the *load broker* (L).

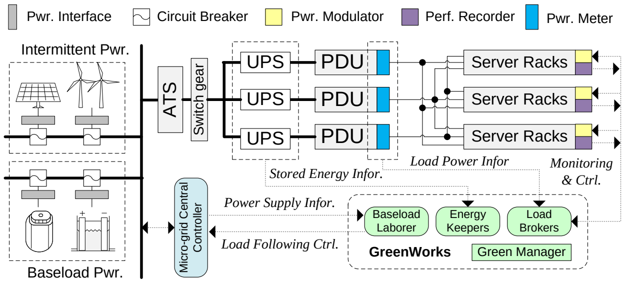


Figure 3: High-level system architecture for GreenWorks.

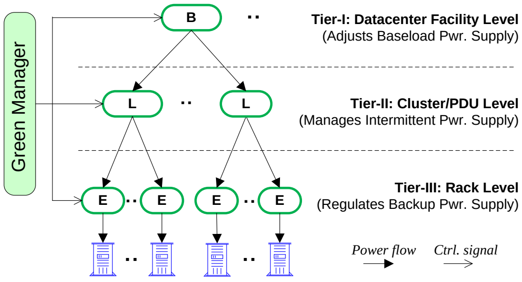


Figure 4: GreenWorks power management hierarchy.

The *baseload laborer* controls the output of distributed generators such as fuel cells and bio-fuel generators. It is responsible for providing a specific amount of baseload power to satisfy the basic power needs (i.e., datacenter idle power). It can also provide load following services [28] at each coarse-grained time interval.

The *energy keeper* is able to provide necessary power support if intermittent power supply drops suddenly or load surge happens. It also monitors the capacity utilization and the health status of the battery packs. In Figure 3, we use distributed battery architecture (at server cluster level) since it has better energy efficiency, reliability, and scalability [31].

The *load broker* is responsible for managing the fine-grained power mismatch between the fluctuating datacenter load and the intermittent power supply. We leverage the performance scaling capability (via CPU frequency scaling) of server system to match load power demand to time-varying green energy budget.

3.2 Power Management Hierarchy

Although the hybrid renewable energy systems are often centrally installed at the datacenter facility level, improving the overall efficiency requires a multi-level, cooperative power management strategy.

GreenWorks uses a three-tier control hierarchy for power management coordination. It organizes different types of green workers in the power management hierarchy based on their design goals.

As Figure 4 shows, in the top tier of the hierarchy is the baseload laborer. We put the load laborer at the datacenter facility level since it is where the baseload power generator is integrated. Managing baseload power budget at datacenter level facilitates load following control, thereby minimizing over-/under- generation of the baseload renewable energy.

GreenWorks manages the intermittent renewable power supply at the cluster level, or PDU (power distribution unit) level. At this level, dynamic voltage and frequency scaling (DVFS) shows impressive peak power management capabilities [32] and could be leveraged to manage the supply-load power mismatch. During runtime, the load broker calculates the total renewable power generation based on the baseload power budget and the assigned renewable power. When the total renewable power generation is not enough, the load broker will decrease server processing speed evenly or request stored energy (from the energy keeper), depending on whichever yields the best design tradeoff.

The energy keeper resides in the third tier of the hierarchy. This allows us to provide backup power directly to server racks if local demand surge happens or power budget drops. Such distributed battery architecture [31] has many advantages such as high efficiency and reliability. In this study, we leverage it for managing fine-grained supply-load power mismatches.

The main advantage of our multi-level cooperative power management scheme is that it facilitates cross-source power optimization. For example, GreenWorks allows datacenters to schedule additional baseload generation to release the burden of the energy backup when the capacity utilization of onsite batteries is high. It also allows them to request additional stored renewable energy to boost server performance if necessary.

4. Multi-Objective Power Management

In this section, we propose multi-source driven multi-objective power management for GreenWorks. The basic idea is to take advantages of the cross-source coordination capability of GreenWorks to balance the usage of different types of energy sources. To achieve this goal, we develop a novel three-stage coordination scheme that synergistically combines battery-aware power management, workload-aware power management, and variability-aware power management to achieve the best design trade-offs.

4.1 Stage I: Adequate Power Supply Budget

The green manager enters power management Stage-I (as shown in Figure 5), when the renewable power generation is unable to ensure the rated speed on all the active servers. In this stage, the excess renewable energy generated will be stored in UPS batteries if there is still enough room. In addition, the green manager also monitors the actual charging current and the maximal power capacity of batteries. The remaining excess renewable power will be send to the utility grid via grid-tie inverter, which is a power inverter that synchronizes onsite power generation with a utility line.

```
Requires:  The percentage of job execution time increase: T
           TimeTable[T][index], a N×2 lookup table for N jobs
Initialize: TimeTable is sorted based on T (descending order)

PowerHeadroom= TotalSupply – PeakLoadDemand;
for each job j in the TimeTable
    if job j has enough thermal headroom then
        while (The frequency of j < maxFreq)
            Increase the node frequency for job j;
            Re-evaluate PowerHeadroom
            if PowerHeadroom = 0  then break;
```

Figure 5: Load adaptation pseudo code for Stage-I.

During runtime, the load broker dynamically monitors each job's progress and calculates an *execution time increase* (ETI). Assuming that a job *j* has *n* execution phases: $\{1, 2, 3, \cdots; i, \cdots; n\}$. For a given execution phase *i* that spends t_a seconds under actual processing frequency f_{actual}, it would spend t_r seconds under rated processing frequency f_{rated}. If we scale down the frequency (i.e., $f_{actual} \leq f_{rated}$), we expect to increase the execution time (i.e., $t_a \geq t_r$). As frequency scaling mainly changes CPU time and has little impact on non-CPU time (i.e., I/O waiting time and memory access time), the job's ETI in phase *i* is given by:

$$T_{ij} = t_a - t_r = \mu(1 - \frac{f_{actual}}{f_{rated}})t_a, \quad \mu = \frac{CPUtime}{Runtime}, \tag{1}$$

where μ is the monitored actual CPU utilization (under scaled processing frequency) in execution phase *i*. ideally, without performance scaling, the total execution time E_r of previous *i* execution phases is:

$$E_r = \sum t_r = E_a - \sum_i T_{ij}, \tag{2}$$

where E_a is the actual total execution time of previous *i* execution phases monitored by load brokers. Thus, we can compute the percentage increase of execution time at the end of execution phase *i* as:

$$T_{\%(\mathrm{i,j})} = \left(\sum_i T_{ij}\right) \Big/ E_r \tag{2}$$

In Figure 5, the green manager dynamically updates the job execution time information and maintains a sorted lookup table for each running job. When allocating additional renewable power budget across server nodes, the green manager will always give priority to jobs that have higher job execution time increase. Specifically, our green manager uses a job acceleration scheme which opportunistically boosts the processing speed/frequency (i.e., over-clocking) to take advantage the additional renewable power budget. This can help mitigate unnecessary energy loss due to power feedback and improve workload performance. It allows a processor to enter a performance state higher than the specified frequency if there is enough thermal/power headroom and if it is enabled by the power management software. Through execution time monitoring and power allocation balancing in the Stage I, we can greatly improve average workload performance.

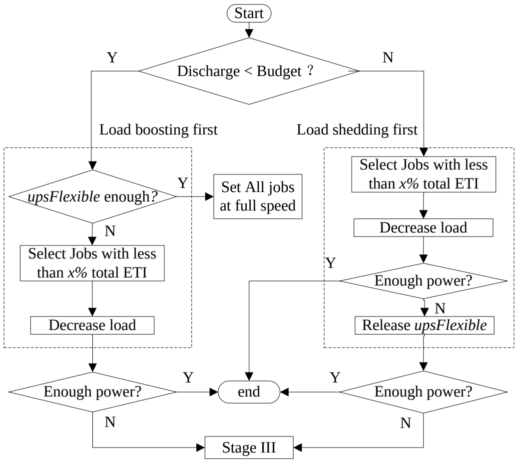


Figure 6: The power management pseudo code for Stage-II.

4.2 Stage II: Moderate Power Supply Drop

Our system enters Stage-II when it senses inadequate power supply. Unlike prior designs which heavily rely on either load shedding or backup power, we use a balanced power management, as shown in Figure 6.

Battery Discharge Control:

Battery lifetime is an important design consideration. To maximize the benefits of the stored energy without compromising reliability, we dynamically monitor the discharge events of the UPS system and calculate a *discharge budget* based on the aggregated discharge throughput (Amp-hours) of the batteries, the overall runtime of the battery, and the rated cycle life.

We use an Ah-Throughput Model [33] to evaluate the battery cycle life and a kinetic battery model (KiBaM) [34] to analyze the battery charging/discharging behaviors. The Ah-Throughput model states that there is a fixed amount of charges that can be cycled through a battery before it requires replacement. The KiBaM model uses a chemical kinetics process as its basis and describes the charge movement inside the battery, as shown in Figure 7. Both models provide reasonable evaluations of battery systems and have been used in professional power system simulation software developed by the National Renewable Energy Lab [35].

We use two different power control schemes in this stage. If the required UPS energy is within the *discharge budget*, the green manager will give priority to using stored energy to maintain high performance (load boosting). Otherwise, it will first decrease the server speed (load shedding) and then use stored energy if necessary. In Figure 7, we assume a maximal UPS discharge amount of 40% of the total installed capacity, which we refer to as *flexible UPS energy* (*upsFlexible*, 0~40% of the total capacity). We also define a *reserved UPS energy* (40%~80% of the total capacity), which is used to handle significant power drop in the Stage-III.

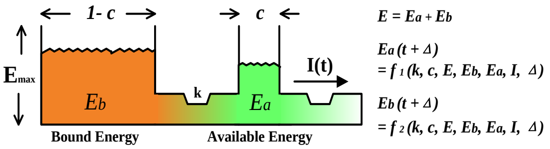


Figure 7: The KiBaM battery model [34]. The stored charge is distributed over two pools: An available-energy pool supplies current directly to the load.

Load Shedding Control:

The load brokers of GreenWorks use performance statistics to make load shedding decisions. GreenWorks allows the datacenter operator to specify a limit (not a hard limit) on job ETI to achieve different performance goals. Our system allows performance scaling only on jobs that have less than *x*% (default value is 10%) increase of execution time. We refer to this as *x% load shedding mechanism*. If there is still a demand-supply mismatch after the x% load shedding and the system has run out of flexible UPS power, the green manager will enter to power management Stage-III.

4.3 Stage III: Significant Power Supply Drop

Our system enters Stage-III when it realizes that moderate load tuning in Stage-II cannot handle the significant power mismatch. The Stage-III is an emergency state since in this scenario the green manager might put the load into minimum power state and use reserved UPS capacity to avoid server shutdown.

Saving UPS Reserved Capacity:

Maintaining an appropriate level of stored energy is important to ensure service availability. In this stage we trade off performance for higher reserved UPS capacity. We first decrease load power demand, and then use stored energy to bridge the remaining power gap.

Deadline-Aware Load Shedding:

GreenWorks uses a deadline-aware load shedding to achieve a better tradeoff between UPS capacity and job execution time increase. Figure 8 shows the algorithm for our deadline-aware load shedding.

The green manager first checks the current ETI values of all the jobs for load shedding opportunities. It calculates a *Time Budget* which evaluates if a job could meet its deadline in the future with frequency boosting techniques. For example, if the monitored CPU utilization μ is 50% (i.e., CPU time is 50% of the job runtime), a 20% frequency increase in the future is expected to reduce 50% × (1-1/1.2) = 0.08s execution time for one second frequency boost.

To estimate the total *Time Budget*, one must know the chances (%) of enabling boosted processing speed. In this study we use historical renewable power traces to estimate the changes of receiving additional renewable power. To further improve accuracy, one can combine our estimation with weather forecasting.

```
Requires: The value of power shortfall after Stage II: Shortfall
Initialize: The mean percentage of CPU time (i.e., utilization): μ
            The duty ratio of performing turbo boost: D
            The likelihood of receiving adequate renewable power: P

// 1st step of Stage III: decrease load power demand
for each job j in the TimeTable
    Saving = μ ×(1 – 1/FreqSpeedup) ;
    TimeBudget = RemainingRuntime × D× P × Saving;
    if the execution time increase of job j < TimeBudget then
        if ( Freq. of j > MinFreq ) & (Shortfall >0) then
            Lower the node frequency for job j;
            Re-evaluate Shortfall;
        if Shortfall < 0 then break;
// 2nd step of Stage III: use reserved UPS energy if have to
if Shortfall > 0 then decrease load in round-robin fashion
Re-evaluate Shortfall;
if Shortfall > 0 then
    if Shortfall < upsReserve then
        release UPS power;
    else shut down servers
```

Figure 8: Power management pseudo code for Stage-III.

Assuming that the given job has 1 hour remaining execution time and the chance of receiving adequate green power is 60%, the anticipated time of being in Stage I is 3600s × 60% = 2160s. However, the actual turbo boost duration is far less than this value. In Figure 5, a duty ratio *D* is defined as the percentage of one period in which the CPU is over-clocked. The value of *D* is hardware-specific and is used to control the thermal headroom of processors. If the duty ratio is 30%, the anticipated turbo boost duration is 3600s × 60% × 30% = 648s. Therefore the total *Time Budget* is 648s × 0.08s/s = 52s. This means that the given job can tolerate up to 52s ETI at the current timestamp.

If the given job has enough *Time Budget*, our controller will incrementally reduce its CPU frequency (Δf=0.1GHz) until it reaches its lowest speed (*MinFreq* = 1.6GHz). It will put server nodes into low power states in a round-robin fashion if the demand-supply discrepancy still exists. Finally, we release the reserved UPS energy if necessary. In this study we assume that each node runs independent data-processing task. Parallel workloads are often not accelerated as much as calculated since the accelerated threads or processes have to wait for others. Exploring workloads with high communication to computation ratio is our future work.

Note that we assume that a job's runtime is known a priori. Typically, HPC users are required to submit their job runtime estimations to enable backfilling, which can help maximize cluster utilization. In this study we leverage it to determine job deadline.

4.4 Managing Baseload at Coarse-Grained Interval

At each fine-grained timestamp (e.g., every 1 second), the green manager adjusts the load processing speed and manages the stored energy. The objective is to mitigate power mismatch caused by the variability issue in the intermittent power supply and server load.

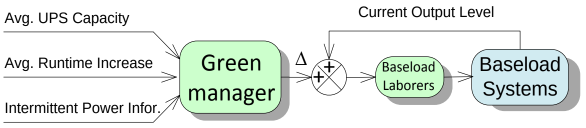


Figure 9: Feedback control for managing baseload power.

At each coarse-grained timestamp (15 minutes), it adjusts the baseload power generation level through the baseload laborers, as shown in Figure 9. The green manager collects the monitored information at the end of each coarse-grained control period; it then adjusts the output of the baseload power supply based on the average power supply shortfall in the last control period. The green manager can also incrementally add additional baseload power (10% of the current output level) if the monitored UPS capacity is low (*upsFlexible* = 0), or the workload performance is low (e.g., 80% of the jobs would be delayed), or the anticipated wind power availability is low (e.g., P > 80% in Figure 8).

5. Evaluation Methodologies

We develop a simulation framework for datacenters powered by renewable generation mix. As shown in Figure 10, this framework is configured into three layers for modeling the entire system from the job dispatching behavior to the power system specifics. It uses discrete-event simulation to process a chronological sequence of job submissions. It also simulates the power behavior of renewable energy system on per-second time scale which is in tune with our datacenter job scheduler. This three-layer framework provides us the flexibility in analyzing various design spaces.

We adopt renewable energy system model from HOMER [35]. Table 2 shows the parameters we used. All the values are carefully selected based on manufacturer's specifications, government publications and industry datasheet. The maximum baseload power output in our simulator equals to the average power demand of the evaluated datacenter workload. The default load following interval is 15 minutes. The capacity of our simulated battery cell is 24Ah at a 20-hour rate (1.2A discharge current). Its capacity is 10Ah at a 15-minute rate (40A discharge current). We determine the total battery capacity in such a way that the backup power system can ensure 15 minutes power output in emergency. We maintain detailed log of each discharging event to calculate battery life using methods in [33, 34].

We use wind turbine as our evaluated intermittent power source since it is widely used to provide abundant and affordable green energy for large-scale facilities. We collect minute-by-minute wind speed data from the National Wind Technology Center [36] during the month of March, 2012, as shown in Figure 11. We calculate wind power based on the wind speed data and the wind turbine output curve.

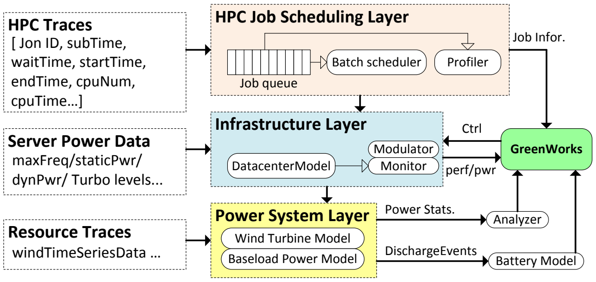


Figure 10: Details of our three-layer simulation platform.

Inputs	Typical Value	Value Used
Load Following Interval	5 min ~ 1 hour	15 min
Battery Life Cycle	5,000 ~ 20,000	10,000 times
Rated Depth of Discharge	0.8	0.8
Battery Efficiency	75% ~ 85%	80%
Max Charging Current	N/A	8 Ah
Peukert Coefficient	1.0~ 1.3	1.2
UPS Installed Backup	10~20 min	15 min

Table 2: Key parameters used in the simulation [21-25].

We use a queueing-based model that takes real workload traces as input. It uses a first come first serve (FCFS) policy and puts each job request into a queue and waits to grant allocation if computing nodes are available. Each job request in the trace has exclusive access to its granted nodes for a bounded duration. Such trace-driven simulation has been adopted by several prior studies on datacenter behaviors and facility-level design effectiveness [8, 19, 37, 38].

We use real-world workload traces from a well-established online repository [39]. As shown in Table 3, these workload activity logs are collected from state-of-the-art HPC systems in production use around the world. We select five key task parameters in each trace file: job arrival time, job start time, job completion time, requested duration, and job size in number of granted CPU nodes. As shown in Table 3, we select eight 1-week workload traces that have different mean utilization level and mean job runtime.

Our datacenter infrastructure is based on the IBM System x3650 M2 (2.93G Intel Xeon X5570 processor) high-performance server which supports Intel Turbo Boost technology. While the number of performance states (P-states) is processor specific, we assume 12 P-states as indicated in [40]. The minimum frequency is 1.6GHz and the normal frequency is 2.9GHz. In Turbo Boost mode, the processor could increase the frequency by 14%. We increase the frequency moderately (i.e., 10%) when the Turbo Boost mode is enabled. Our power model uses CPU utilization as the main signal of machine-level activity. Prior work has shown that CPU utilization traces can provide fairly accurate server-level power prediction [41]. According to the published SPEC power data, the modeled system consumes 244 Watts at full utilization and 76 Watts when idle [42].

Traces	**Descriptions**	**Load**	**Mean Inter-arrival**		**Avg. Job Run Time**	
Thunder	Lawrence Livermore Lab's 4096-CPU capacity cluster called Thunder	61%	1.8 min	Short	0.58h	Short
DataStar	San Diego Supercomputer Center's 184-node cluster DataStar	56%	3.5 min		1.41h	
Atlas	Lawrence Livermore Lab's 9216-CPU capability cluster called Atlas	33%	11 min	Long	0.61 h	
BlueGene	A 40-rack large Blue Gene/P system at Argonne National Lab	26%	8.4 min		1.4h	
RICC	A massively parallel Japanese cluster of cluster with 1024 nodes	49%	0.9 min	Short	16.6 h	Long
MetaC	Czech national grid infrastructure called MetaCentrum	67%	2.1 min		11.8 h	
Seth	A 120-node European production system named Seth	80%	21 min	Long	6.2 h	
iDataPlex	320-node IBM iDataPlex cluster for Climate Impact Research	18%	50 min		3.7h	

Table 3: The evaluated real-world workload traces in representative HPC datacenters [39].

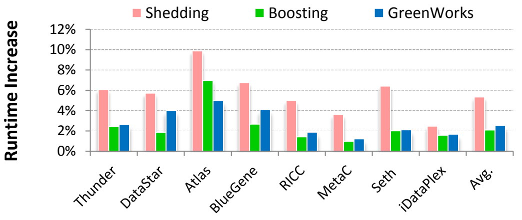


Figure 12: Average increase of job turnaround time (i.e., the average ETI for all the processed jobs).

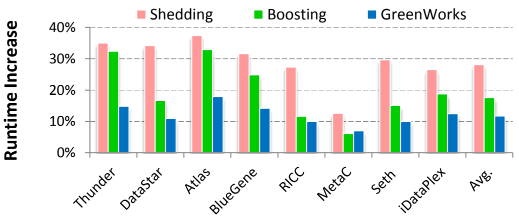


Figure 13: Maximum increase of job turnaround time (i.e., the average ETI for the worst 5% delayed jobs).

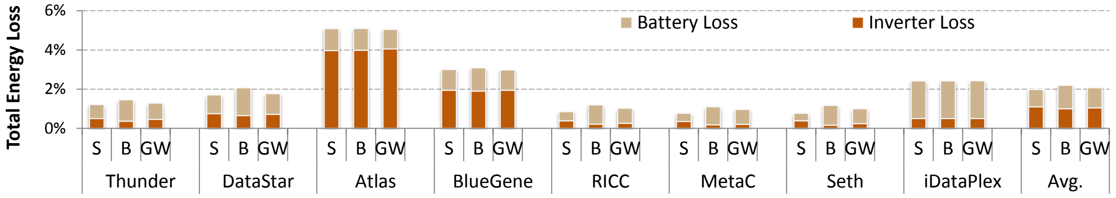


Figure 14: GreenWorks (GW) maintains almost the same green energy utilization efficiency as *Shedding* (S) and *Boosting* (B).

6. Results

In this section we evaluate the benefits of applying GreenWorks to datacenters powered by hybrid onsite green power supplies. We compare GreenWorks to two state-of-the-art baselines: *Shedding* and *Boosting*. *Shedding* is a widely used load management schemes for emerging renewable energy powered datacenters [43, 44]; *Boosting* represents recent datacenter power management approaches that emphasis the role of energy storage devices [45, 46]. Both baselines use UPS and server load scaling to manage fine-grained power shortfall and adjust baseload output level at each end of the control period. The only difference between the two is that *Shedding* gives priority to load scaling, while *Boosting* gives priority to UPS stored energy.

6.1 Execution Time

We evaluate datacenter performance in terms of average job turnaround time increase compared to an oracle (which always ensures full processing speed with zero service downtime). Figure 12 shows the average job execution time increase. On average, the job execution time increase of *Shedding*, *Boosting* and GreenWorks are 5.4%, 2.1%, and 2.4%, respectively. Compared to *Shedding*, *Boosting* shows less execution time increase since it trades off UPS capacity for performance. As GreenWorks seeks a balanced power management across different power supplies, it yields slightly higher ETI compared to *Boosting*.

The performance of the worst 5% jobs could significantly affect the service-level agreements (SLA) of datacenters. Figure 13 shows the maximum increase of job turnaround time which is calculated as the average execution time increase of the 5% worst cases. The worst-case result of *Shedding* is 28%. Surprisingly, GreenWorks (12%) reduces the maximum job execution time increase by 33%, compared to *Boosting* (18%). The improvement is due to the *x% shedding mechanism* (detailed in Section 4.2). By modifying the value of the *x*, one can easily adjust the performance goal of GreenWorks (detailed in Section 6.5).

6.2 Energy Efficiency

The main sources of inefficiency in green datacenters are the battery round-trip power loss and the power conversion loss in the grid-tied inverter. We assume a typical battery system of 80% round-trip energy efficiency and a power inverter of 92% energy efficiency.

GreenWorks could maintain the same energy efficiency as *Shedding* and *Boosting*. In Figure 14 we show the total energy loss due to the battery round-trip energy loss and the inverter's power conversion loss. The overall efficiencies of the three evaluated schemes are very close to each other. The differences are less than 0.5%. Compared to the other two, *Boosting* shows relatively lower inverter loss because it can maximally leverage the power smoothing effect of UPS battery to reduce the amount of power feedback.

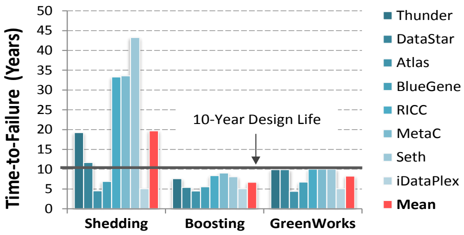


Figure 15: The estimated battery lifetime based on detailed battery charging/discharging statistics.

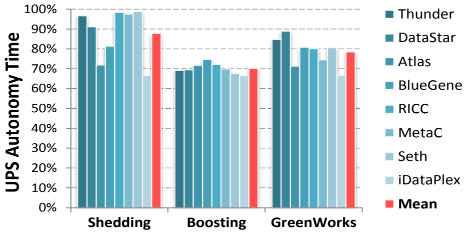


Figure 16: The normalized backup time throughout the evaluated operation duration (normalized to rated backup time).

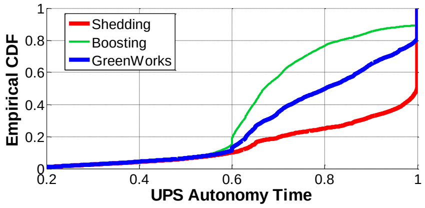


Figure 17: Cumulative distribution function (CDF) for the normalized UPS autonomy time.

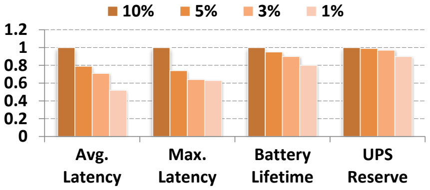


Figure 18: Sensitivity to various performance capping requirements. The default performance threshold is 10%.

6.3 Battery Lifetime

Typically the rated lifetime of a valve-regulated lead–acid battery (VLRA) is 3 years to 10 years [47]. In Figure 15, GreenWorks shows a near-threshold battery life (8.3 years). It means our multi-source multi-objective power management can maximally leverage batteries without degrading their life significantly. In contrast, *Boosting* shows a mean lifetime of 6.7 years; and *Shedding* shows a mean lifetime of 19.7 years. Typically, the battery lifetime is not likely to exceed 10 years [47]. The reason *Shedding* over-estimates battery life is that the system underutilizes batteries. Since batteries may fail due to various aging problems and self-discharging issues, it is better to fully utilize it.

6.4 UPS Backup Time

Another advantage of GreenWorks is that it can optimize the mean UPS autonomy time. The autonomy time is also known as backup time. It is a measure of the time for which the UPS system will support the critical load during an unexpected power failure. Figure 16 shows the mean normalized UPS autonomy time throughout the operation duration for various datacenter traces and different power management schemes. On average, the mean autonomy time is: *Shedding* (88%), *Boosting* (70%), and GreenWorks (78%).

In Figure 17 we plot the cumulative distribution function (CDF) for the normalized UPS autonomy time. Our results show that the CDF curve of GreenWorks lays nicely between our two baselines: *Shedding* and *Boosting*. GreenWorks could ensure rated backup time (the discharge time of a fully charged UPS) for 20% of the time. *Shedding* maintains its rated backup time for 50% of the time and the number for *Boosting* is only 10%. This is because *Boosting* uses UPS battery much more aggressively than *Shedding*.

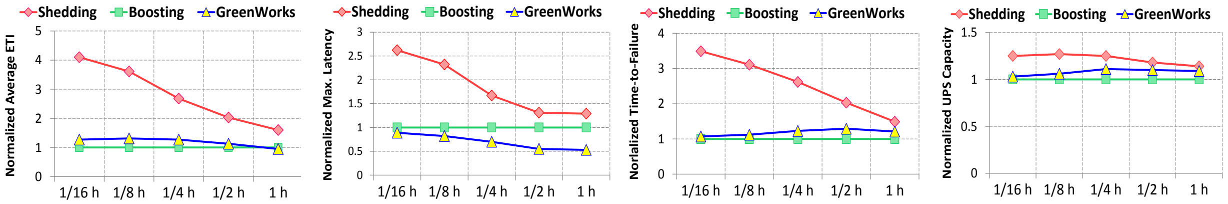
Energy storage devices should be always taken care of. A lower autonomy time can pose significant risk as the backup generator may not be ready to pick up the load. Without appropriate power management and coordination, datacenters have to increase their installed UPS capacity, which is both costly and not sustainable.

6.5 Control Sensitivity

We also evaluate the control sensitivity of our system by varying the value of several key parameters.

In Figure 18 we first show the impact of the *x% shedding mechanism* (detailed in Section 4.B) on various performance metrics of GreenWorks. The default value of the performance limit in our study is 10% and we evaluate the performance impact when the user lowers the threshold. As can be seen, the *x% shedding mechanism* has a much larger impact on the average latency, other than the battery lifetime and UPS capacity. Decreasing the threshold (i.e., the *x*) can reduce the job execution time and increase the reliance on energy storage elements, which will lower the battery lifetime and backup capacity to some extent.

In Figure 19 we further evaluate the impact of the control intervals (load following intervals of the baseload power supply) on the performance of our multi-source driven multi-objective control. Our default interval of adjusting the baseload power is 15 minutes. All the results are normalized to that of *Boosting*. They show that the job latency drops as the control interval becomes larger. The battery lifetime and UPS capacity of GreenWorks both rise as we increase the length of the control interval. Note that although the relative latency may decrease as load following interval increases, the actual value of latency increases. A longer interval often degrades load following effectiveness, and therefore increases the chance of power mismatch.



(a) Job ETI (b) Worst-case ETI (c) Battery time-to-failure (d) UPS reserved capacity

Figure 19: Sensitivity to various load following intervals of the baseload power supply.

7. Related Work

Managing computer systems on green energy has been done at various levels. However, existing designs mainly focus on certain specific type of green energy sources (i.e., intermittent power or baseload generators) and overlook the benefits of cross-source coordination.

Managing Intermittent Power Source: Prior studies on this issue mainly focus on load adaptation schemes which can be broadly categorized into three types: load shedding, load deferring, and load migration. For example, SolarCore [6] is a load shedding based design. It temporarily lowers server power demand using per-core power gating when solar power drops. [48] and [49] investigate server power adaptation under intermittent power budget. Load deferring, also known as load shifting, leverages the flexibility of job scheduling [9, 18, 50]. It re-schedules load by shifting user requests to a future time horizon if renewable power is currently not available. Load migration based design focuses on re-allocating application to another datacenter that has reserved capacity [8]. With intelligent workload packing and virtual machine placement [51], one could further minimize resource wastage and power consumption in green datacenters.

Managing Baseload Power Generation: Several recent proposals have explored baseload power supply in datacenters. The most similar studies are [52] and [19]. In [52], the authors propose design methodology for sustainable datacenters powered by onsite generation. However, they mainly focus on high-level datacenter infrastructure management policies. In [19], we investigate the benefits of load following mechanism in distributed generation powered datacenters. However, [19] does not consider the power variability issue of intermittent green power integration.

Managing Backup Power / UPS Systems: There have been several studies exploring the use of backup power systems for energy-efficient datacenters. For example [31, 45, 46, 53] investigate the use of energy storage (particularly the UPS system) to manage the datacenter peak power. For example, [31] explore the TCO of the distributed UPS system in datacenters and propose using local distributed UPS to shave the datacenter peak power. Govindan et.al, [53] use UPS as the major tuning knob for minimizing power cost in aggressively under-provisioned datacenter infrastructure.

Cost-Aware Green Energy Scheduling: The system cost-effectiveness also receives many attentions in renewable energy powered datacenter. For example, [10] proposes algorithms that minimize fossil fuel-based energy consumptions; [11] discusses load balancing on distributed datacenters. Recent work in capacity planning for datacenters also looks at the cost issue of green energy purchases [3].

In contrast to prior work, this paper explores hierarchical, cross-layer power management for datacenters powered by hybrid renewable energy systems. We consider an integrated mix of complementary power provisioning methods that include intermittent power supply, baseload power generation, and energy storage devices.

8. Conclusions

In this paper we investigate green datacenters powered by a mix of various green energy sources and energy storage devices. Although these emerging power systems are often centrally installed at the datacenter level, maximizing the overall datacenter efficiency requires a hierarchical power management strategy. We propose GreenWorks, a novel framework that could greatly facilitate multi-source based green datacenter design. GreenWorks enables datacenters to make informed power management decisions based on the available baseload power output, renewable power variability, battery capacity, and job performance. We show that GreenWorks could achieve less than 3% job runtime increase, extend battery life by 23%, increase UPS backup time by 12%, and still maintain desired overall energy efficiency.

Acknowledgement

This work is supported in part by NSF grants 1320100, 1117261, 0937869, 0916384, 0845721 (CAREER), 0834288, 0811611, 0720476, by SRC grants 2008-HJ-1798, 2007-RJ-1651G, by Microsoft Research Trustworthy Computing, Safe and Scalable Multi-core Computing Awards, by NASA/Florida Space Grant Consortium FSREGP Award 16296041-Y4, by three IBM Faculty Awards, and by NSFC grant 61128004. Chao Li is also supported by University of Florida Graduate Fellowship, Yahoo! Key Scientific Challenges Program Award, and Facebook Fellowship. Jingling Yuan is supported by NSFC grant 61303029.

References

[1] DCD Industry Census 2011: Forecasting Energy Demand, 2011, www.dcd-intelligence.com

[2] Data Center Carbon Calculator, http://www.apcmedia.com/salestools/WTOL-7DJLN9_ R0_EN.swf

[3] C. Ren, D. Wang, B. Urgaonkar, and A. Sivasubramaniam, Carbon-Aware Energy Capacity Planning for Datacenters, *IEEE Int'l Symp. on Modeling, Analysis & Simulation of Computer and Telecommunication Systems*, 2012

[4] N. Deng, C. Stewart, D. Gmach, M. Arlitt, and J. Kelley, Adaptive Green Hosting, *Int'l Conf. on Autonomic Computing*, 2012

[5] M. Haque, K. Le, I. Goiri, R. Bianchini, and T. Nguyen, Providing Green SLAs in High Performance Computing clouds. *Int'l Green Computing Conference* , 2013

[6] C. Li, W. Zhang, C. Cho, and T. Li, SolarCore: Solar Energy Driven Multi-core Architecture Power Management, *IEEE Int'l Symp. on High-Performance Computer Architecture* , 2011

[7] N. Deng, C. Stewart, J. Kelley, D. Gmach, and M. Arlitt, Adaptive Green Hosting, *Int'l Conf. on Autonomic Computing* , 2012

[8] C. Li, A. Qouneh, and T. Li, iSwitch: Coordinating and Optimizing Renewable Energy Powered Server Clusters, *Int'l Symp. on Computer Architecture* , 2012

[9] I. Goiri, K. Le, T. Nguyen, J. Guitart, J. Torres, and R. Bianchini, GreenHadoop: Leveraging Green Energy in Data-Processing Frameworks, *ACM EuroSys*, 2012

[10] K. Le, R. Bianchini, M. Martonosi, and T. D. Nguyen, Capping the Brown Energy Consumption of Internet Services at Low Cost," *Int'l Green Computing Conference*, 2010

[11] Z. Liu, M. Lin, A. Wierman, S. Low and L. Andrew, Greening geographical load balancing, *ACM Int'l Conf.e on Modeling and Measurement of Computer Systems*, 2011

[12] http://biomassmagazine.com/articles/8351/microsoft-data-center-to-install-biogas-fuel-cell-power-plant

[13] http://www.datacenterknowledge.com/archives/2012/05/30/hp-developing-net-zero-data-center-concept/

[14] http://green.ebay.com/greenteam/ebay/blog/Building-a-Greener-Company/26

[15] http://www.apple.com/environment/renewable-energy/

[16] Enabling the Low Carbon Economy in the Information Age, http://www.smart2020.org

[17] G. Burch, Hybrid Renewable Energy Systems, Natural Gas / Renewable Energy Workshops, U.S. Department of Energy, 2001

[18] I. Goiri, W. Katsak, K. Le, T.D. Nguyen, and R. Bianchini, Parasol and GreenSwitch: Managing Datacenters Powered by Renewable Energy, *Int'l Conf. on Architectural Support for Programming Languages and Operating System*s, 2013

[19] C. Li, R. Zhou, and T. Li, Enabling Distributed Generation Powered Sustainable High-Performance Data Center, *IEEE Int'l Symp. on High-Performance Computer Architecture* , 2013

[20] R. Lasseter and P. Piagi, Microgrid: A Conceptual Solution, *IEEE Annual Power Electronics Specialists Conference* , 2004

[21] D. Salomonsson, L. Soder, and A. Sannino, An Adaptive Control System for a DC Microgrid for Data Centers, *IEEE Transactions on Industry Applications*, pp. 1910 – 1917, Volume:44, Issue:6

[22] R. Wang, N. Kandasamy, C. Nwankpa, and D. Kaeli, Datacenters as Controllable Load Resources in the Electricity Market, *Int'l Conf. on Distributed Computing Systems*, 2013

[23] W. Deng, F. Liu, H. Jin, and C. Wu, SmartDPSS: Cost-Minimizing Multi-source Power Supply for Datacenters with Arbitrary Demand, *Int'l Conf. on Distributed Computing Systems*, 2013

[24] B. Kirby and E. Hirst, Customer-Specific Metrics For the Regulation and Load-following Ancillary Services, Technical report, ORNL, 2000

[25] The Role of Distributed Generation and Combined Heat and Power (CHP) Systems in Data Centers, Technical Report, US EPA, 2007

[26] S. Chowdhury and P. Crossley, Microgrid and active distribution networks, The Institute of Engineering and Technology, 2009

[27] Fuel Cell Technologies Program Multi-year Research, Development and Demonstration Plan, Technical Report, US Department of Energy

[28] H. Zareipour, K. Bhattacharya, and C. Canizares, Distributed generation: current status and challenges, *the 36th Annual North American Power Symposium*, 2004

[29] The Importance of Flexible Electricity Supply, Solar Integration Series, Technical Report, U.S. Department of Energy, 2011

[30] National Survey on Data Center Outages, Ponemon Institute, White Paper, 2010

[31] 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 Data Centers, *Int'l Symp. on Computer Architecture* , 2012

[32] X. Fan, W. Weber, and L. Barroso, Power Provisioning for a Warehouse-Sized Computer, *Int'l Symp. on Computer Architecture* , 2007

[33] H. Bindner, T. Cronin, P. Lundsager, J. Manwell, U. Abdulwahid, and I. Gould, Lifetime Modelling of Lead Acid Batteries, Technical Report, Risø National Laboratory, 2005

[34] M. Jongerden and B. Haverkort, Which Battery Model to Use?, the 24th UK Performance Engineering Workshop, 2008

[35] Getting started guide for HOMER version 2.1, National Renewable Energy Laboratory, 2005

[36] National Wind Technology Center (NWTC), http://www.nrel.gov/wind/

[37] S. Pelley, D. Meisner, P. Zandevakili, T. Wenisch and J. Underwood, Power Routing: Dynamic Power Provisioning in the Data Center, *Int'l Conf. on Architectural Support for Programming Languages and Operating Systems*, 2010

[38] F. Ahmad and T. Vijaykumar, Joint Optimization of Idle and Cooling Power in Data Centers While Maintaining Response Time, *Int'l Conf. on Architectural Support for Programming Languages and Operating Systems*, 2010

[39] Logs of Real Parallel Workloads, http://www.cs.huji.ac.il/labs/parallel/workload

[40] Host Power Management in VMware vSphere 5, Technical Report, VMware, 2010

[41] P. Ranganathan, P. Leech, D. Irwin, and J. Chase, Ensemble-level Power Management for Dense Blade Servers, *Int'l Symp. on Computer Architecture*, 2006

[42] SPEC power_ssj2008 http://www.spec.org/power_ssj2008/

[43] G. Ghatikar, V. Ganti, N. Matson, M. Piette, Demand Response Opportunities and Enabling Technologies for Data Centers: Findings from Field Studies, Technical Report, Lawrence Berkeley National Laboratory, 2012

[44] H. Xu, U. Topcu, S. Low, C. Clarke, and K. Chandy, Load-shedding Probabilities with Hybrid Renewable Power Generation and Energy Storage, *the 48th Annual Allerton Conference on Communication, Control, and Computing*, 2010

[45] D. Wang, C. Ren, A. Sivasubramaniam, B. Urgaonkar, and H. Fathy, Energy Storage in Datacenters: What, Where, and How much?, *ACM Int'l Conf. on Modeling and Measurement of Computer Systems*, 2012

[46] S Govindan A. Sivasubramaniam and B. Urgaonkar , Benefits and Limitations of Tapping into Stored Energy for Datacenters, *Int'l Symp. on Computer Architecture*, 2011

[47] S. McCluer, Battery Technology for Data Centers and Network Rooms: Lead-Acid Battery Options, APC White Paper #30

[48] C. Li, R. Wang, N. Goswami, X. Li, T. Li, and D. Qian, Chameleon: Adapting Throughput Server to Time-Varying Green Power Budget Using Online Learning, *Int'l Symp. on Low Power Electronics and Design*, 2013

[49] C. Li, Y. Hu, R. Zhou, M. Liu, L. Liu, J. Yuan, and T. Li, Enabling Datacenter Servers to Scale Out Economically and Sustainably, *Int'l Symp. on Microarchitecture*, 2013

[50] I. Goiri, R. Beauchea, K. Le, T.D. Nguyen, M. Haque, J. Gui-tart, J.Torres, and R. Bianchini, GreenSlot: Scheduling Energy Consumption in Green Datacenters, *Int'l Conf. for High Performance Computing, Networking, Storage and Analysis*, 2011

[51] J. Xu, and J. Fortes, A Multi-Objective Approach to Virtual Machine Management in Datacenters, *Int'l Conf. on Autonomic Computing*, 2011

[52] P. Banerjee, C. Patel, C. Bash, and P. Ranganathan, Sustainable Data Centers: Enabled by Supply and Demand Side Management, *Design Automation Conference* , 2009

[53] S. Govindan，D. Wang， A. Sivasubramaniam, and B. Urgaonkar, Leveraging Stored Energy for Handling Power Emergencies in Aggressively Provisioned Datacenters, Battery Emergency, *Int'l Conf. on Architectural Support for Programming Languages and Operating Systems*, 2012