Oasis: Scaling Out Datacenter Sustainably and Economically

Chao Li, *Member, IEEE*, Yang Hu, Juncheng Gu, Jingling Yuan, and Tao Li

Abstract—As big data applications proliferate, datacenters today are increasingly looking to adopt a scale-out model. Nevertheless, power capacity has become an important bottleneck that restricts horizontal scaling of servers, especially in datacenters that oversubscribe power infrastructure. When a datacenter hits its ceiling for power provisioning, conventionally the owner has to either build another facility or upgrade existing infrastructure—both approaches add huge cost, require significant time, and can further increase carbon footprint. This paper proposes Oasis, a novel datacenter expansion strategy that enables power-/carbon- constrained servers to scale out economically and sustainably. The basic structure of Oasis, called Oasis Node, naturally supports incremental capacity expansion with near-zero environmental impact since it leverages modular solar panels and distributed battery systems to power newly added servers. To optimize the operation of newly added nodes, we further propose a management framework called Ozone. It allows Oasis to jointly perform power supply switching and server speed scaling to improve efficiency locally and globally. We implement a prototype of Oasis and use it as a research platform for evaluating the design tradeoffs of green scale-out datacenters. With Oasis, a green datacenter could gradually double its capacity with near-oracle performance, extended battery lifetime, and 26 percent cost savings.

Index Terms—Data center, power management, scalabilty, energy efficiency

1 Introduction

The server industry is in the midst of a major expansion period spurred by the heightening demand for cloud computing and big data analytics. To support business growth, many datacenters are continually adding computing resources (i.e., scaling out) to their existing sites. During the past five years, this trend has led to significant growth in the number of data center servers. It has been shown that the global server market size is projected to double every five years, accounting for over 1,400 TWh annual energy consumption in 2020 [1].

Over time, power has become an increasingly critical driver for datacenter planning. The constant influx of computing resources in datacenters could eventually become power-constrained. According to two major industry surveys [2], [3], nearly one third of the cooperate datacenter managers expected to run out of power capacity within 12 months. In Fig. 1 we show several widely adopted solutions to the power capacity shortage problem. Most companies turn to server consolidation to avoid server sprawl (i.e., too many under-utilized servers) and free up power capacity. However, without enough power budget, even consolidated servers have to limit their performance using either software-based (e.g., virtual CPU allocation) or hardware-based (e.g., frequency scaling) control knobs to avoid the costly circuit-breaker tripping. While leasing collocated server racks and purchasing third-party cloud services fit small companies on a budget, they are not suitable for large-scale enterprise datacenters that need to sustainably improving computing capability and capacity.

Re-sizing the infrastructure can be a radical solution that allows one to add more servers, racks, and even containerized clusters. However, like building a new datacenter, upgrading initial power system can be a great undertaking. Conventional datacenters often rely on a centralized multi-layer power delivery path that involves many power equipments of different sizes (Fig. 2). Upgrading power infrastructure requires re-design of the entire power delivery path which is not only costly but also time-consuming. Worse, in many urban areas the utility power feeds are often at their capacity and additional electricity access for datacenter is restricted.

Building a scale-out datacenter is challenging since servers are not only power-constrained, but also carbon-constrained. Conventional data centers mainly rely on fossil fuel based utility power. Without green computing techniques and carbon-aware designs, the associated data center carbon emission could reach 100 million metric tons per year worldwide [1]. In recent years, integrating green energy into datacenters has shown great promise in IT carbon capping [4], [5]. This approach has grown from a whisper to a major theme for many vendors seeking to secure energy availability, lowering environmental impact, and improving brand reputation. Several companies, including Microsoft, IBM, Google, and Facebook, all start to explore green datacenter design. For example, Apple's Maiden datacenter in

- *C. Li is with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China. E-mail: chaol@sjtu.edu.cn.*
- *Y. Hu and T. Li are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611. E-mail: huyang.ece@ufl.edu, taoli@ece.ufl.edu.*
- *J. Gu is with the Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, MI 48109. E-mail: gujuncheng@ufl.edu.*
- *J. Yuan is with the School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430070, China. E-mail: yjl@whut.edu.cn.*

Manuscript received 20 May 2016; revised 27 Aug. 2016; accepted 25 Sept. 2016. Date of publication 6 Oct. 2016; date of current version 14 June 2017. Recommended for acceptance by B. Urgaonkar.
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/TPDS.2016.2615625

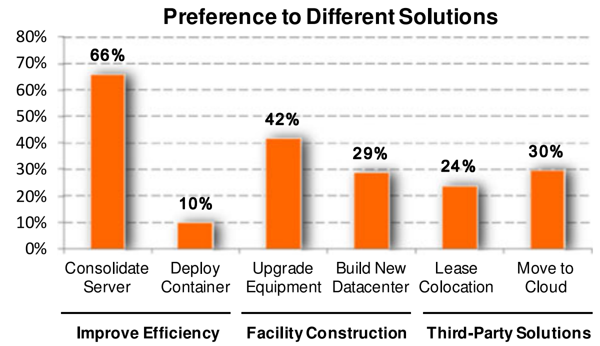


Fig. 1. Three types of solutions to the power capacity problem. Consolidating servers and upgrading power systems are currently the most popular approaches. [3].

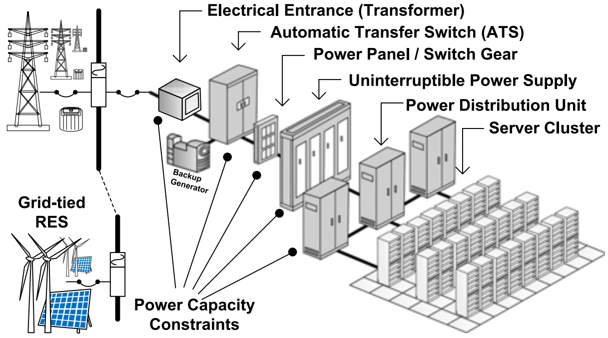


Fig. 2. Conventional datacenter power distribution system is not scale-out friendly. Facility-level green energy integration can also be cost-prohibitive and low-efficiency.

North Carolina draws renewable power from both self-generation and a regional plant.

Unfortunately, the scalability of existing green energy integration schemes is not satisfactory. Most datacenters employ centralized power integration architecture and largely overlook the modularity of renewable power supplies. As shown in Fig. 2, one of the advantages of such facility-level integration is that the renewable power supplies can be synchronized to the utility grid. Although grid-tied renewable power supply eliminate power variability problem, their operation heavily rely on the availability of utility power. This centralized integration not only creates a potential single-point of failure, but also makes future capacity expansion expensive.

In this study we ask one question: *"can we enable a power-/carbon- constrained datacenter to sustainably scale out with the minimum upfront cost?"* Faced with the growing computing demand, the skyrocketing energy cost, and the looming environmental crisis, this question is very important to IT companies who have to embrace new power provisioning and management schemes to survive.

We propose *Oasis*, a sustainable and economical capacity expansion strategy for scale-out datacenters. Its basic structure is called *Oasis Node*, which consists of server resources, power systems, and a control hub. Different from prior works, this specific architecture allows datacenters to add computing resources along with power capacity increments. In addition, *Oasis* takes advantage of modular renewable energy systems and distributed battery systems to ensure that every Watt added is green and scalable. Therefore, a datacenter can grow with smaller carbon footprint and less power capping needs using *Oasis*. More importantly, *Oasis* means a way to significantly reduce the upfront investment that conventional centralized system design may have.

To fully unleash the potential of *Oasis*, we further propose *Ozone*, a management framework for *Optimized Oasis Operation (O3)*. Transcending the boundaries of primary and secondary power supplies in conventional designs, *Ozone* provides a power source switching mechanism that enables server system to dynamically identify and select the most suitable power sources (i.e., utility power and solar power in this work). It could distill crucial runtime statistics of different power sources to avoid inefficient power drawn or aggressive battery discharge and in the meantime adaptively adjust server speed to improve workload performance.

This paper describes our best practices to ensure that the *Oasis* design is a benefit and works to its full potential. We have implemented the *Oasis Node* as a prototype. Its current form is a micro server rack (12U) that draws power from solar panels, utility power, and local energy storage devices. We have made *Oasis* a unified system that synergistically integrates energy source switching, power system monitoring, and architectural support for power-aware computing. Our 1st generation system (managed by *Ozone Lite*) mainly emphasizes the automatic control of local nodes through a power control hub (PCH) built from scratch. Our 2nd generation design (managed by *Ozone Pro*) makes one step further to provide cluster-level power management and global coordination.

Overall, *Oasis* and *Ozone* constitute an interesting research platform from many perspectives: (1) it enables datacenter power capacity to scale out along with server resource expansion; (2) it links power supply and server system to enable joint management of energy source and compute workload; (3) its power provisioning architecture (i.e., hybrid energy source + distributed control domain) brings improved availability; (4) it provides the flexibility of offering customized green IT services based on different user expectations. The main contributions of this work can be summarized as follows.

- We propose *Oasis*, a novel on-demand datacenter expansion strategy that enables server resources and the associated power system to jointly scale out. It greatly facilitates datacenter capacity planning and green energy integration.
- We propose *Ozone*, an optimization framework for managing green scale-out datacenters. It can extend battery lifetime by 50 percent, increase backup time by 1.9X, while it still maintains near-oracle performance and high green energy usage rate.
- We build a system prototype from scratch. The system offers smart interaction between server system and hybrid power supply, thereby enabling power-driven server load control and application-driven energy source management.
- We analyze the cost benefit of *Oasis* and *Ozone* under large-scale deployment. We show that it enables datacenters to double their capacity with 29 percent less capital expenditure (CapEx) and could reduce 26 percent operating cost (OpEx).

2 Overview of Oasis Architecture

A growing number of data centers are nearing their capacity today. In these datacenters, servers are unable to scale out due to several key constraints such as power

TABLE 1
Comparison of Different Power Capacity Scale-Out Models

	Carbon Footprint	Capacity Scalability	Utility Power Cost	Utility Energy Cost	Allow for Green Power?	Green Power Cost
Utility Power Over-Provisioning	High	Poor	High	High	Yes	High
Centralized Green Power Expansion	Reduced	Improved	Reduced	Reduced	Yes	Reduced
Distributed Green Power Increment	Low	Good	Reduced	Reduced	Yes	Low

demand, carbon footprint, floor space, and cost. In this work we focus on constraints that are closely related to power and energy. This section first compares typical power expansion strategies which we refer to as capacity scale-out models. We then introduce two key design features of the *Oasis* green computing architecture.

2.1 Scale-Out Models

We classify existing power capacity scale-out model as either utility power over-provisioning or centralized power expansion. In Table 1 we qualitatively compare the scale-out model of *Oasis* with existing strategies.

The utility power over-provisioning is the most convenient and straightforward model. The datacenter power delivery infrastructure in this case is designed to support the maximal utility power the load may ever draw in the expected service life. Although abundant power headroom has been provided for future resource expansion, this approach results in under-utilized power systems and requires significant upfront cost. Installing onsite renewable energy systems could further exacerbate the low utilization problem. Therefore this model normally chooses utility power as the primary energy source and the carbon footprint can be very high.

The second scale-out model supports most of the prior green datacenter designs [4], [5]. In this case, the power delivery infrastructure is provisioned for a certain level of anticipated future power drawn and the power capacity expansion is handled by datacenter-level renewable power integration. Although this model improves datacenter efficiency and economy in certain degree, it still outgrows the capacity of renewable power systems.

Oasis explores a "pay-as-you-grow" model for scale-out datacenters, which we refer to as distributed power increments. The utility power configuration of *Oasis* is similar to the centralized power expansion approach: power systems are provisioned for a fixed level of load power demand to avoid over-provisioning. However, when datacenter reaches its maximum capacity, we add renewable power capacity by small increments in a distributed manner. This approach not only provides carbon-free power capacity expansion to a power-constrained datacenter, but also avoids over-committing capitals in the renewable energy systems.

2.2 Modularized Systems

Oasis is built upon various modularized systems. Its basic structure is termed as *Oasis Node*. As shown in Fig. 3, each *Oasis Node* consists of a group of green energy connected racks and a couple of related power systems. Each group of racks are attached to a distributed power control hub, which further connects to the utility power distribution unit (PDU), onsite renewable power supply, and local batteries. While pre-fabricated racks and PDUs have long been equipped in datacenters, the use of modular renewable energy systems and energy storage systems are still emerging elements in datacenter design.

2.2.1 *Solar Panels with Micro-Inverters*

Renewable energy systems are ideal power supplies that support carbon-free power capacity expansion. They are usually modular and highly scalable in their capacity.

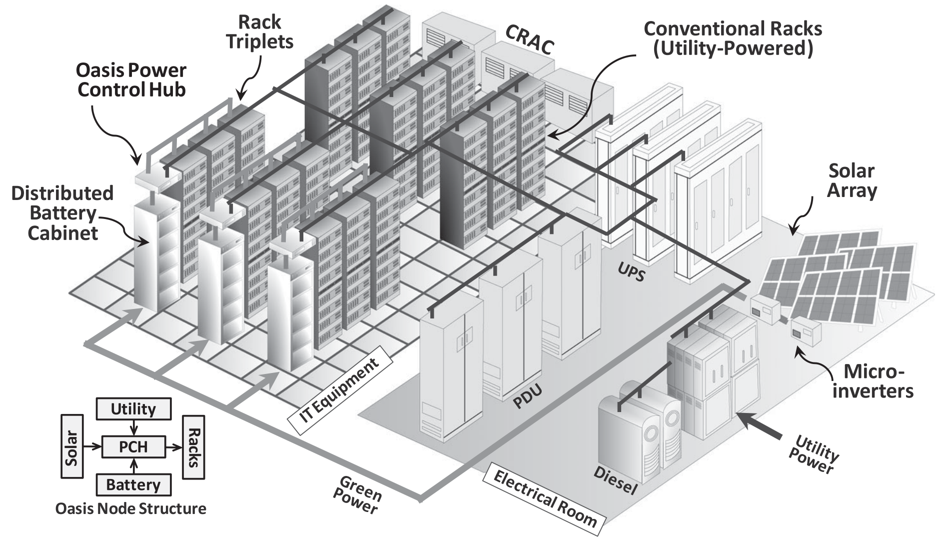


Fig. 3. Schematics of Oasis architecture. Oasis leverages the modularity of both battery systems and renewable power systems to improve scalability and economy. A power control hub is designed to coordinate utility power, solar power, and local battery.

Compared to wind turbines, solar panels can provide even smaller capacity increments, especially when equipped with micro-inverters [6].

Conventionally, solar power system uses bulky string inverters which require several panels to be linked in series to feed one inverter. String inverters are prone to failure from heat and shows low efficiency. In contrast, a micro-inverter based panel has its own inverter. Micro-inverters are much smaller in size and can be built as an integrated part of the panel itself. This gives solar power systems better scalability, reliability and efficiency.

2.2.2 Distributed Energy Storage Devices

We leverage distributed energy storage system to incrementally add backup power as computing demand increases. Google and Facebook first employ distributed energy storage for reducing the energy efficiency loss. Such de-centralized design also avoids a single-point of failure (SPOF) and increases availability. Recently, distributed battery has been used to shave peak power [7]. In scale-out datacenters, they are necessary components for storing renewable energy and handling power outage.

Our design adopts the distributed battery provisioning architectures in the Open Compute Project [8]. As shown in Fig. 3, a battery cabinet populated with commodity lead-acid batteries is used to provide standby power for a rack triplet. The battery cabinet includes a high current DC bus bar, breakers, fuse disconnects, and sensors. In Facebook's original design, the battery carbinet could maintain around 45 seconds at full load.

2.3 Distributed Integration

A distinctive feature of *Oasis* is that it integrates renewable power at the server rack level. As shown in Fig. 3, this integration method allows us to scale out server system on a per-rack basis. We do not use centralized power integration since it does not support fine-grained server expansion very well. Today's power-constrained datacenter typically over-subscribes pieces of its power distribution hierarchy such as the power distribution unit, thereby creating power delivery bottlenecks. Adding renewable power capacity at datacenter-level power switch gear does not guarantee increased power budget for the newly added server racks since probably the associated PDU has already reached its capacity limit.

Oasis chooses to route renewable power to each server racks using power source switching other than power synchronization. In the power system domain, synchronization refers to as the minimization of the variances in voltage, phase angle, and frequency between the renewable power supply and the utility grid power so that the two power sources can work together. *Oasis* does not synchronize renewable power for three considerations:

1) *Efficiency: Oasis* integrates renewable power system at the rack level, but synchronization is normally done at the facility level. At the datacenter facility-level, grid-tied renewable power system can cause additional efficiency loss. The newly added renewable power incurs many levels of power redundant conversion before reaching server racks, resulting over 10 percent energy efficiency loss [9].
2) *Reliability:* Many issues and challenges remain to be considered for successful synchronization [10]. Due to the time-varying nature of renewable power, synchronization may induce voltage transients and frequency distortions that can jeopardize system stability. As the impact of power quality problems on server reliability is still an open question, adding grid-tied renewable power at the rack level is debatable.
3) *Availability:* As required by UL Std. 1741 and IEEE Std. 1547, all grid-tied inverters must disconnect from the grid if they detect power islanding. That is, these renewable energy systems must shut down if the grid power no longer present. Today the U.S. datacenter experiences 3.5 times of utility power loss per year with an average duration over 1.5 hours [11]. With synchronization based design, datacenters may lose renewable power when they need it most.

In this study we aim to provide non-intrusive power integration that allows datacenters to increase power capacity on-demand and reduce carbon footprint gradually. Section 3 propose a novel power-/workload- cooperative management scheme for scale-out green datacenter and Section 4 shows our implementation of *Oasis*.

3 Managing Green Scale-Out Datacenters

Managing the newly added green computing racks and the associated power capacity increments can be challenging. This section proposes *Ozone*, a power management scheme for *Optimized Oasis Operation* (O3). *Ozone* enables *Oasis* users to take a huge step forward towards building efficient and scalable green computing environment.

Built upon *Oasis*, *Ozone* features a set of control policies for improving efficiency and reducing cost. Among many factors that may affect *Oasis* operation, five are most important. Three of them are power-related: *utility budget*, *solar budget*, and *load demand*. Two of them are battery-related: *discharge budget* and *remaining capacity*. In this section we first introduces power source switching, the basic control of *Ozone*. We then discuss the local and global optimization that *Ozone* can provide.

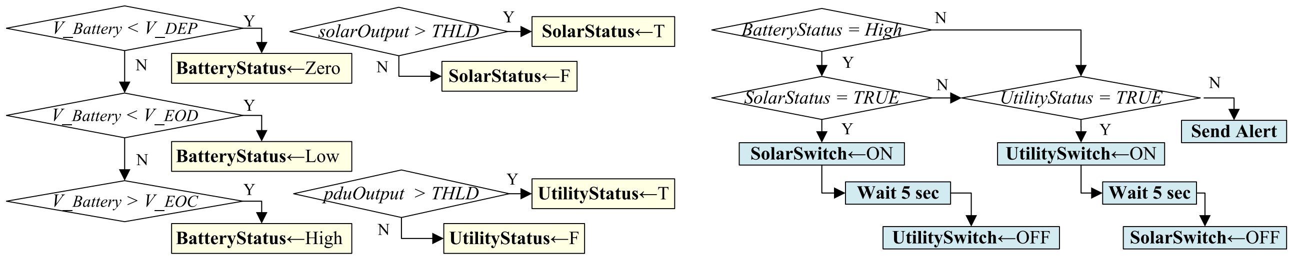
3.1 Power Source Switching

The first priority of *Ozone* is to ensure reliable power provisioning for *Oasis nodes*. *Oasis* allows each node to dynamically select its own power supply through power source switching offered by our PCH unit. To make data center more reliable and efficient, *Ozone* supports two kinds of power source switching modes: 1) autonomous switching and 2) coordinated switching.

3.1.1 Autonomous Switching

In this mode, *Oasis Nodes* run autonomously, i.e., switching server between solar panel and utility grid based on the monitored solar and utility power budget. To achieve this, the PLC in our power control hub defines two modules: SupplySense and SupplySwitch, as shown in Fig. 4.

The *SupplySense* module focuses on setting parame-ters that are used for making power source switching decisions. For example, the server rack is allowed to discharge battery only when the battery voltage exceeds a pre-defined end-of-charge voltage (V_EOC). The system starts to charge the



(a) *SupplySense* module (b) *SupplySwitch* module

Fig. 4. The control flow of power supply switching. Oasis in autonomous mode uses two atomic modules to control the system. (DEP - Depletion; EOD - End of Discharge; EOC - End of Charge; V_EOC> V_EOD> V_DEP).

battery only when its terminal voltage drops below a predefined end-of-discharge voltage (V_EOD). The charge/discharge thresholds prevent batteries from entering deep-discharging or over-charging status. To avoid power failure, our system double-checks the solar and utility power to see if they have reached the minimum allowed output values.

The *SupplySwitch* module executes power source switch with reliability support. For every switch operation, the controller first checks the output of power supplies to ensure that they work normally. In Fig. 4, the PCH configuration allows an overlap between solar power supply and utility power supply when performing power supply switching. We disconnect one energy source only when the other has been successfully working for a short time (5 seconds). This helps to avoid potential switching failure.

The autonomous mode is the default mode for each newly added *Oasis* node and it is important for two reasons. First, it makes *Oasis* node a plug-and-play unit. Different nodes can work independently without the needs for complex configuration or user intervention in resolving resource conflicts. Second, it improves power switching reliability. No matter how renewable power generation output varies, the *Oasis* can always secure appropriate power supply that the server rack needs.

3.1.2 Coordinated Switching

Oasis also provides servers the option of establishing their own power supply switching policies. Our system currently allows two user-defined operations: Utility Power Enforcement and Solar Power Enforcement. Users could specify their preferred energy source at runtime, by calling the power management agent. However, the execution of a power supply switching signal depends on battery status, solar power output, and utility power availability. The user's switching request will be ignored if it violates the power budget or causes safety issues.

The coordinated switching is useful for managing multiple *Oasis Nodes*. As more green energy powered servers are added, *Oasis Nodes* must relate to each other in a meaningful way to achieve the best design tradeoff.

3.2 Local Power Management

Ozone seeks a balanced control between power supply switching and server load management on each *Oasis Node*. Its goal is to maintain a desired trade-off between carbon footprint, battery lifetime and workload performance. During runtime, our hardware agents could inform *Ozone* the power supply behaviors, the health status of batteries, as well as workload performance.

3.2.1 Discharge Credit

Batteries made from electrochemical cells typically have a lifespan (e.g., 5~years). They become no longer suitable for storing energy in mission-critical data centers after the designed service life.

Ozone uses a discharge credit to control battery usage. Here the discharge credit refers to how much amp-hour energy can be discharge from the battery. *Ozone* uses the Ah-Throughput Model [12] for calculating the discharge credit. This model assumes that there is a fixed amount of aggregated energy (total discharge throughput) that can be cycled through a battery before it requires replacement. During runtime, *Ozone* monitors battery discharging events and calculates battery throughput (in amp-hour) based on Peukets's equation [13]:

$$Discharge = I_{\text{actual}} \cdot (I_{\text{actual}}/I_{\text{nominal}})^{pc-1} \cdot t. \tag{1}$$

In Equation (1), I_{actual} is the observed discharging current, $I_{nominal}$ is the nominal discharging current given by the manufacturer, *pc* is the Peukets coefficient, and *t* is the discharging duration. Over time, the aggregated discharge increases. To avoid battery over-utilization, *Ozone* sets a limit on battery usage. Assuming D_{total} is the overall discharge credit, the discharge credit is set as:

$$D_B = D_{\text{total}} \times T/T_{total} - \sum_i Discharge_i, \tag{2}$$

where T is the time of usage (to date) of the battery and T_{total} is the expected total battery service time. At the beginning of each control cycle, the *Oasis Nodes* receive a discharge credit which specifies the maximum amount of stored energy discharge that will not compromise battery lifetime. The discharge credit affects both power supply switching and server workload control. When it is inadequate, *Ozone* prefers to switch the server back to utility power (with increased carbon footprint) or decrease power demand to avoid heavily utilizing batteries.

Blindly capping battery usage can cause under-utilization of the battery and reduced cost-efficiency. Enforcing strict discharge credit greatly degrades workload performance since for every monitoring interval T_c, the discharge credit is fixed and limited ($D \times T_c/T_{total}$). Any unexpected load surge can quickly use up the precious discharge credit and cause expensive performance capping or power switching to the carbon-intensive utility grid.

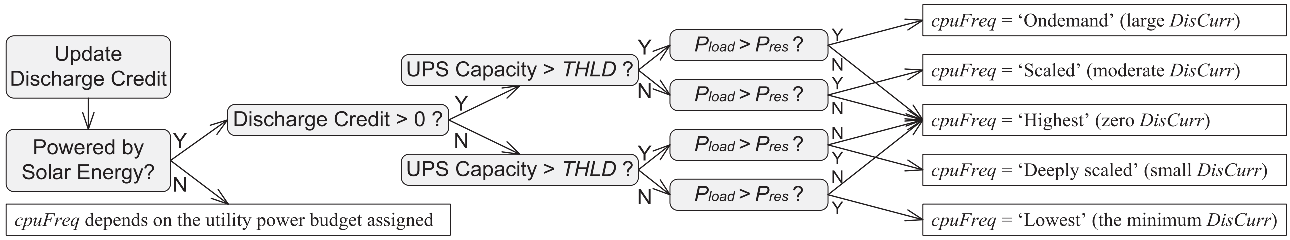


Fig. 5. The decision tree of Ozone performance management. THLD: Preset threshold of the state of charge; cpuFreq: Server frequency mode; DisCurr: Battery discharge current.

Different from its Lite version, *Ozone Pro* uses a relaxed discharge credit control strategy. If the current discharge credit is zero and the job performance degradation is untolerable (e.g., >2 percent), *Ozone Pro* will add the monitored energy throughput of the last control period to the credit. This enables *Oasis Node* to opportunistically use future discharge credits to handle transient load surge and inadequate power budget. When job traffic is low or the renewable power is high, the battery units are rarely used. In this period we can regain discharge credit and pay the debt.

3.2.2 Performance Tuning

Ozone adaptively adjusts server load based on power supply states along with battery management. Our system supports five coarse-grained performance levels: Highest, Ondemand, Scaled, Deeply Scaled, and Lowest.

In Fig. 5, the server speed not only depends on discharge credit, but also depends on battery capacity. Maintaining necessary backup capacity is critical since batteries not only smooth out time-varying renewable power, but also serve as uninterruptible power supplies for scale-out servers. The backup capacity is the primary factor that determines UPS autonomy time (a.k.a. backup time). It is a measure of the time for which the UPS system will support the critical load during utility power failure.

Ozone sets a limit on the minimum remaining capacity of batteries. Our system only uses 40 percent of the installed capacity for managing renewable power shortfall (referred to as flexible capacity). The remaining 60 percent battery capacity (i.e., reserved capacity) is mainly used for emergency handling purpose. In Fig. 5, for example, when the load power demand is lower than the solar power demand, we always run the server at its highest speed. When both discharge credit and flexible capacity are adequate, *Ozone* gives high priority to server performance boost (i.e., at the highest frequency) with the support of battery. When the system runs out of flexible capacity but still have discharge credit, *Ozone* allows the server to keep using green energy with reduced server frequency. If the discharge credit is not enough, the server may run at the "Deeply scaled" or "Lowest" performance levels, depending on the state of the charge of the batteries.

3.3 Global Power Management

Another important issue is how multiple *Oasis Nodes* are coordinated. If the installed *Oasis Nodes* keep running at the autonomous mode, it is very likely that at certain point all the loads are switched to the utility grid side. This can cause utility power surge, which significantly increases power cost. Fig. 6 shows our coordinated control scenario and Fig. 7 shows the basic algorithm.

Cluster-level power management is provided with *Ozone Pro* for alleviating the burden on the utility grid. The idea is to limiting the number of utility-connected nodes without causing service disruption. *Ozone Pro* puts all the nodes into a charging queue and manages the queue length (i.e., the number of utility-connected nodes). In Fig. 6, the utility-connected nodes are waiting to be switched to the solar power side. They are sorted based on their batteries' state of charge (SOC) and their batteries are charged by solar panels. On the solar panel side, nodes are in discharging mode. Once a node has used up its stored energy, *Ozone Pro* will trigger a node swap. It first switches high-SOC nodes (at the front of the queue) to the solar power supply; then it switches the node with depleted battery to the utility grid.

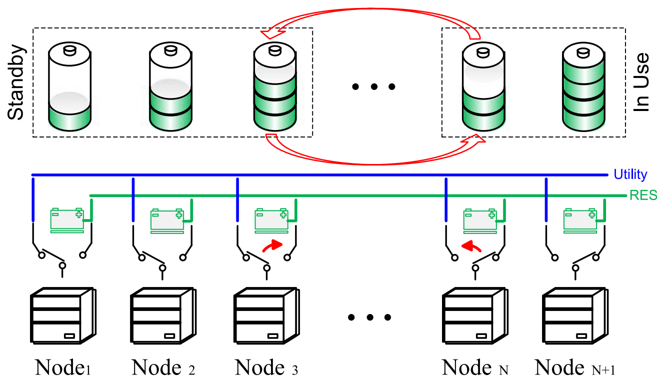
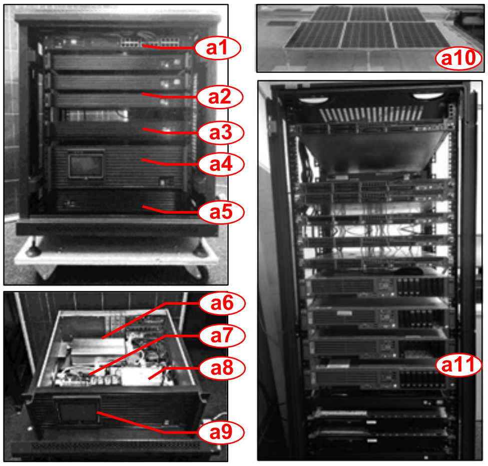


Fig. 6. The node swapping scenario of Oasis.

Definition:
SwitchStatus - switch connection status (1: solar; 0: utility);
SOC - a battery's state of the charge (0%~100%);
N_s - number of standby node; N - number of total nodes

for each load monitoring timestamp *T*
  $c \leftarrow$ average charging time
  $d \leftarrow$ average discharge time
  $r \leftarrow$ ceil(c/d)
for each load tuning timestamp *tick*
  for each *Oasis* Node *i*
    if (SwitchStatus[*i*] == 1) && (SOC[*i*] < 20%)
      swapID $\leftarrow$ standby battery with the largest SOC
        if ($N_s > N \times r/(r+1)$)
        switch *swapID* to solar system first
        switch *i* to utility grid
      else switch *i* to utility grid
    if (SwitchStatus[*i*] == 0) && (SOC[*i*] > 95%)
      switch *i* to solar system
  Update SwitchStatus

Fig. 7. The node swapping algorithm of Ozone.



a1) network switch, **a2**) computing nodes, **a3**) front-end server, **a4**) power control hub, **a5**) In-rack battery chassis, **a6**) converter and inverter, **a7**) power supply switch, **a8**) PLC module, **a9**) HMI, **a10**) roof-mounted solar panels, **a11**) local cloud servers

Fig. 8. Prototype implementation of Oasis.

Ozone Pro intends to avoid simultaneously switching all the nodes from the solar power side to the utility power side. To dynamically maintain a stable charging queue, the number of utility powered nodes (N_U) and the number of solar-powered nodes (N_S) needs to roughly follow:

$$N_U/N_S = T_{\text{charging}}/T_{\text{discharging}}. \tag{3}$$

In Equation (3), $T_{charging}$ is the average charging time and $T_{discharging}$ is the observed current discharging time. For example, assume there are ten nodes and the charging time is 1 hour and the observed current discharging time is 15 minutes, then N_U is set as 8 and N_S is set as 2. If the workload varies greatly or the renewable power output changes, *Ozone* needs to adjust the queue accordingly.

4 Prototype Implementation

We build a prototype of *Oasis Node* from scratch and setup our management framework. Fig. 8 shows the major components of our implementation and Table 2 shows the basic configuration of *Oasis* and *Ozone*.

4.1 Computing Equipment

Our scaled-down *Oasis Node* is composed of four 1U rack-mounted servers. They are high-performance low-power nodes that use Intel Core i7-2720QM 4-core CPU as the processing engine. The measured idle power and peak power of each server are 21 and 55 W, respectively. With the Intel Turbo Boost Technology, these processors support up to 3.3 GHz operating frequency.

We deploy Xen 4.1.2 with Linux kernel 2.6.32.40 on each server node. Both para-virtualization and hardware virtualization are used to support different virtual machines with various memory size. Multiple virtual machines are booted to execute different workloads on each server. We enable the relocation feature of VM in Xen and VM live migration. Xen power management feature is also enabled to dynamically tune the vCPU frequency. As default, our system kernel is configured with the on-demand frequency scaling governor. We set the minimum frequency as 0.8 GHz and the normal frequency as 2.2 GHz.

TABLE 2
Computing Platform Configuration

Oasis Computing Node	
M/B	SuperMicro ITX Socket G2
CPU	Intel Core i7-2720QM, 4-core, 2.2 G, 45 W
Memory	8 GB registered
Storage	Seagate Barracuda 7200RPM, 500 GB
O/S	Xen 4.1.2 with Linux kernel 2.6.32.40

Oasis Front-end Server	
M/B	SuperMicro ATX Socket C32
CPU	AMD Opteron 4256EE, 8-core 2.5 G, 32 W
Memory	16 GB registered Memory
Storage	Seagate Barracuda 7200RPM, 1000 GB
Switch	TP-Link TL-SF1024 24-Port, 10/100M

Ozone Central Manager	
Server	HP ProLiant DL360 G5 compute server
CPU	2 dual-core Xeon 3.2 GHz, 16 G RAM
NAS	4 HP ProLiant DL380 G5 storage servers
Disk	1TB SAS 15000RPM HDD
Switch	Cisco Linksys SRW2024

4.2 External Power System

We build a solar system consisting of six roof-mounted solar panels. It could generate 7 kWh AC power a day on average under local weather condition in May in the State of Florida, US. This external power supply is enough to power both *Oasis Node* and servers that run *Ozone*.

The solar power output is usually a complex function of the solar irradiation, ambient temperature, and the load connected to panels. To harvest solar energy, we have customized a stack of nine 2Ah sealed lead-acid batteries for each *Oasis Node* and charge them using a maximal power point tracker (MPPT). Our battery chassis can provide 5~30 minutes of backup time, depending on the server load and solar energy availability. The MPPT samples the output of the solar cells and applies appropriate control to maximize the solar generation and optimize the battery usage under varying solar irradiation conditions.

4.3 Computing Equipment

The power control hub is the major hardware addition of the *Oasis Node*. It is a small system that consumes negligible power (about 9 Watts). It integrates battery charger, power inverter, switch panel, micro-controller, and interface that allows easy system diagnosis.

The basic function of the PCH is to switch the server between renewable power and utility power. Fig. 9 shows a schematic diagram inside of the control hub. The solar system provides DC solar power to feed the PCH. We first convert DC solar power to AC to match the output level of utility power distribution unit. The AC solar power and AC utility power are merged (but isolated) at a switch panel inside of the PCH. We use high voltage Omron relay to perform power switch and a Mitsubishi FX2N programmable logic controller (PLC) is used to manage the switch behavior. Finally, the PCH routes power to rack-level power distribution strip

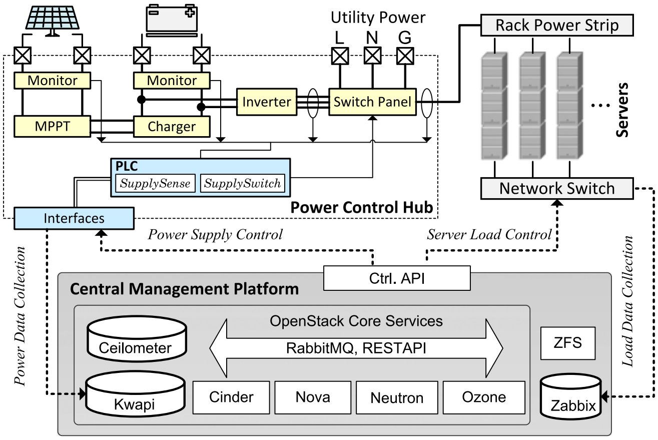


Fig. 9. System architecture of Oasis.

which further feeds a cluster of servers. The entire rack is either powered by the utility power or renewable power, depending on the status of the internal power switch.

Another feature of the PCH is that it sets up the communication gateway between the power supply layer and the server workload layer. This design is partly enlightened by the energy management strategy of the smart grid, which focuses on intelligent communication and control across electrical loads, power electric interfaces, and generators. Each PCH features a human-machine interface (HMI). It is a touch screen panel with build-in microprocessors that could display graphics and interchange data to and from the PLC. The HMI allows external system to communicate the PCH using Modbus, a widely used serial communication protocol for industrial electronic devices due to its robustness and simplicity. Typical Modbus TCP communication includes a Modbus server and a Modbus client. In this work, our central power management platform is the client (master). It initiates requests periodically through socket to the server (slave), i.e., the HMI device.

Finally, we enrich each PCH with various sensing components that keep monitoring real-time solar power output, battery voltage, and discharge current. This systematic checkup of the power supply offers a real-time profile of the system's energy utilization. All the data are exposed to the central management platform (Fig. 9).

4.4 Management Platform

We setup four HP ProLiant DL360 G5 rack-mounted servers as a central cloud datacenter. The total power demand of the compute cluster rages between 1.1 and 1.8 KW. Each server has two low-voltage dual-core Xeon 5100-series 3.2 GHz processors and 16 GB RAM. They are connected through a Cisco Gigabit switch and have access to our network-attached storage (NAS) system (four HP ProLiant DL380 G5 storage servers).

We deploy OpenStack cloud management suite to build our production IaaS environment on the compute cluster. We use OpenStack Cinder on the NAS system to provide elastic storage service for computing instances. In order to enable inline deduplication we deploy ZFS and use ZVOL to provide volume service for Cinder.

We implement *Ozone* as a software module in 8K LOC of Python and C. It is a hierarchical framework that oversees

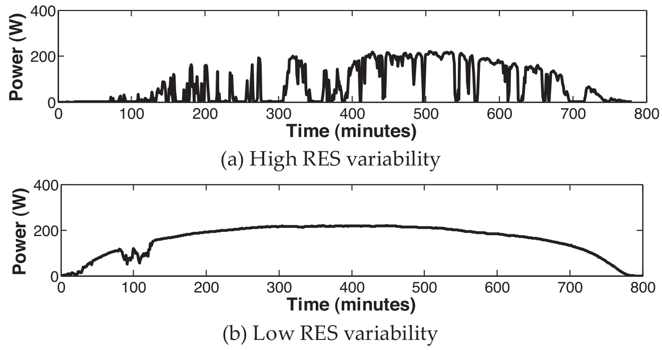


Fig. 10. Evaluated real world renewable energy supply scenarios.

each *Oasis Node*. We deploy different modules to collect data from the power infrastructure level to cloud middleware level. These modules store runtime information using MySQL database and expose RESTful API for easy access of resources from external services.

At the power system level, we leverage OpenStack Kwapi to collect readings from various devices. These devices normally use different communication protocols such as Modbus and SNMP. We design drivers to encapsulate APIs for solar systems and intelligent PDUs.

At the compute resource level, we choose enterprise-level monitoring software Zabbix to collect data from the cloud hardware. We monitor runtime performance information such as CPU utilization, network bandwidth, disk bandwidth and RAM of each server nodes.

At the cloud middleware level we leverage OpenStack Telemetry (Ceilometer) service as our cloud monitoring tool. Ceilometer was first developed for cloud metering in terms of CPU and network costs. Now it is capable to collect various metrics not only with the goal of metering but also monitoring and alarming. We use Ceilometer to capture the events of cloud components such as compute service, network service and storage service.

5 Evaluation Methodology

We develop an evaluation framework for *Oasis*. This framework combines prototype-based experiment with trace-based simulation. We test various workloads and power management schemes on our prototype under different renewable energy conditions. The trace-based simulation is mainly used for comparing different power management schemes under different solar power supply traces. In addition, since we implement only one *Oasis* node, the trace-based simulation is also used to evaluate the peak power optimization effectiveness at a larger scale (a ten-node *Oasis* cluster). Overall, the framework is configured into two layers: the *Power Control Layer*, the *Oasis Operation Layer*, and the *Data Analytic Layer*.

In the Power Control Layer, we feed the system with predefined power budget. We use the peak server power demand as default utility power budget of our system. To ensure fare comparison, we pick up two most representative traces from our system and use it as renewable power budget for all the experiments. In Fig. 10a, the power trace collected in a cloudy day has a peak-to-average ratio of 2.94, which means very high power variability. In Fig. 10b, the power trace collected in a sunny day has a peak-to-average

TABLE 3
Evaluated Data Center Workloads

Abbr.	Workload	Category
Rank	Page rank algorithm of Mahout	Web Search
Nutch	Apache Nutch indexing	Web Search
Bayes	Bayesian classification	Machine Learning
KM	K-means clustering	Machine Learning
Web	Web serving	Internet Service
Media	Cline-server media streaming	Internet Service
YCSB	Yahoo! cloud serving benchmark	Cloud Application
SWtest	Software testing	Cloud Application

TABLE 4
Evaluated Management Schemes

Schemes	Description
Oasis-B	Battery-oriented design that focuses on opportunistically discharging batteries
Oasis-L	Load-oriented design that use frequency scaling to minimize battery discharge
Ozone Lite	Balanced usage of server load adaptation and stored renewable energy
Ozone Pro	Global optimization with peak power shaving and better battery management

ratio of 1.45 (nearly half of the first trace), which means low power variability and high output.

In the *Oasis* Operation Layer, we choose various data center workloads from Hibench [14] and CloudSuite [15]. Hibench consists of a set of representative Hadoop programs including both synthetic micro-benchmarks and real-world applications. CloudSuite is a benchmark suite designed for emerging scale-out applications that are gaining popularity in today's datacenters. In Table 3, we select eight workloads from four categories. Within each experiment, a workload is executed iteratively.

In the Data Analytic Layer, we deploy front-end network server to communicate with the server cluster through a TP-Link 10/100 M rack-mounted switch. The network server uses an AMD low power 8-core CPU with 16 GB memory. We write system drivers using Linux socket to enable data communication between front-end server and the *Oasis Node*. We use Watts UP Pro power meter to collect instantaneous power consumption with high accuracy ($\pm$1.5 percent) and store history power data. We store the collected battery usage and the measured server power data in a log file. We assume the battery has a cycle life of 5,000 times and a maximal service life of ten years.

6 Results

In this section we evaluate the impact of various power supply switching and server adaptation schemes on green scale-out data centers. Specifically, we evaluate three kinds of *Oasis* power management schemes: *Oasis-B*, *Oasis-L*, and *Ozone*. We also compare *Ozone Pro* with *Ozone Lite* to show the benefits of a more intelligent control. Table 4 summarizes the features of different schemes.

6.1 Job Performance

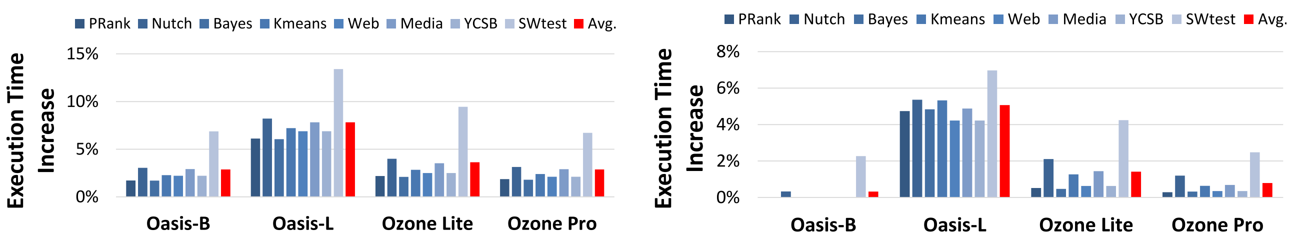
Job turn-around time is an important metrics for data-analytic workload. To evaluate the impact of performance scaling on job turn-around time, we define a lower-is-better metrics called execution time increase (ETI). It represents the difference between the ideal job turn-around time (without performance scaling and service interruption) and the actual job turn-around time.

Fig. 11 presents the ETI of different workloads under both high-variability and low-variability solar power generation scenarios. Our results show that *Oasis-B* yields the best performance and *Oasis-L* has the worst performance. This is because *Oasis-B* trades off battery lifetime for server performance while *Oasis-L* acts contrarily. It is also clear that *Ozone* shows less performance degradation. The results of both *Ozone Lite* and *Ozone Pro* are very close to *Oasis-B* which heavily uses battery to provide power shortfall. By comparing the ETI of *Ozone Lite* with *Ozone Pro*, we can see that *Ozone Pro* could further reduce performance degradation. On average, the ETI of *Oasis-L*, *Oasis-B*, *Ozone Lite*, and *Ozone Pro* is 1.5, 6.4, 2.5, and 1.8 percent respectively. *Ozone Pro* shows near-oracle performance because it allows some critical jobs to use additional discharge budget and utility power when there is a transient load surge or power shortfall.

6.2 Battery Lifetime

Longer battery lifetime (the total service time before replacement) is often favored as it lowers the total cost of ownership (TCO). In Fig. 12 we estimate the battery lifetime based on detailed battery usage log of each cycle.

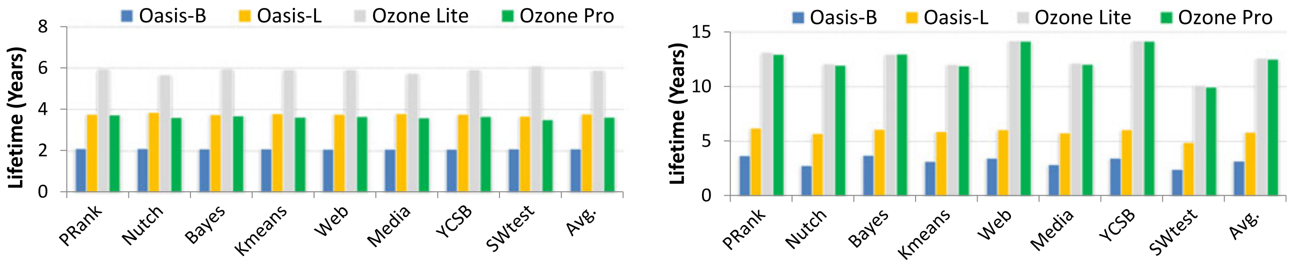
When renewable power output varies significantly, the operation of server nodes typically requires substantial support from the energy storage. As a result, the anticipated battery lifetime is much shorter than the designated service life (5$\sim$10 years), as shown in Fig. 12a. On average, the lifetime of *Oasis-B*, *Oasis-L*, and *Ozone Lite* is 2.1, 3.8, and 5.9 years, respectively. Due to the over-use of battery systems, the service life of *Oasis-B* is only 55 percent of *Oasis-L*. In the figure, *Ozone Lite* could increase battery lifetime by over two years since it actively switch the load to the utility grid to avoid battery deep discharging.



(a) High RES Variability

(b) Low RES Variability

Fig. 11. The execution time increase due to server performance scaling.



(a) High RES Variability (b) Low RES Variability

Fig. 12. The estimated battery lifetime calculated from charge/discharge logs.

When renewable power supply becomes adequate and stable, as shown in Fig. 12b, the lifetime results of all the three schemes increase. This is consistent with our expectations that batteries in this specific scenario are not discharged for most of the time. However, commercial batteries cannot last for over ten years even if they are rarely utilized. Many other issues such as aging and self-leakage can become the dominant factors that prevent batteries from operating for extended duration. In fact, in the real world batteries are discharged much more frequently since the solar panel does not maintain peak output throughout a year. Therefore the actual battery lifetime of *Ozone Lite* should be longer than 6 years but shorter than 12 years.

The results of *Ozone Pro* also depend on the renewable power availability. If renewable power output is low, *Ozone Pro* shows an average lifetime of 3.5 years which is very close to *Oasis-L*. This is because *Ozone Pro* uses a relaxed battery discharge policy for better performance. However, if the renewable power output is high, *Ozone Pro* does not have to aggressively use the discharge credit. As a result, the difference between *Ozone Pro* and *Ozone Lite* is very small. Overall, *Ozone Pro* is able to provide a more attractive performance-lifetime tradeoff (Figs. 11 and 12). It allows data centers to further improve workload performance, while it only moderately increases battery usage frequency.

6.3 Emergency Handling

A low battery backup capacity can pose significant risk since the backup generator may not be ready to pick up the compute load. Fig. 13 shows the average backup capacity for different power management mechanisms. *Ozone* maintains around 73 percent backup capacity when renewable power variability is high and about 98 percent backup capacity when renewable power variability is low. *Oasis* suffers increased numbers of charge/discharge cycles in circumstances that renewable power varies significantly. Therefore, the backup capacity is low for all the three power management schemes in Fig. 12a. Without setting a limit on battery usage and the minimum stored energy levels (i.e., *Oasis-B*), the backup time can drop by 40~70 percent.

6.4 Energy Usage Profile

Leveraging green energy to provide additional power could save utility power bills considerably and lower the negative environmental impact of carbon-constrained datacenters. In Fig. 14 we evaluate green energy utilization as the ratio of renewable energy usage to overall IT energy consumption. While *Ozone* yields impressive system performance, battery lifetime, and battery backup capacity, it shows relatively lower green energy utilization. Compared to *Oasis-B*, *Ozone Lite* yields 19 percent less renewable power rate when renewable power varies significantly and 7 percent less renewable power rate when renewable power variation is low. The reason *Oasis-B* shows high green energy usage is that it heavily uses battery to harvest renewable energy.

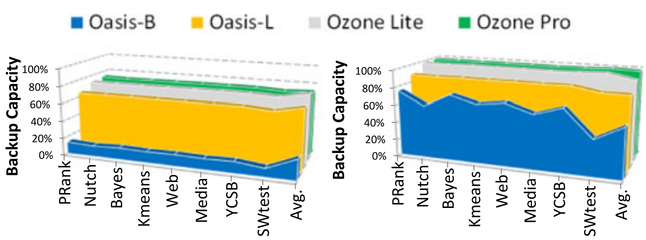
Note that the renewable power variability can greatly affect the green energy usage ratio. When the renewable power budget varies significantly, our results show that *Ozone Pro* yields higher green energy usage compared to *Ozone Lite*, as shown in Fig. 14a. This is because *Ozone Pro* encourages renewable energy harvesting. It does not aggressively scales down load power demand to match the green power budget variation.

7 Cost Analysis

Ozone Pro not only reduces the operating cost (OpEx) of datacenters, but also avoids the significant upfront capital investment. In this section we estimate the cost benefits of *Oasis* and *Ozone* based on our system protptoye.

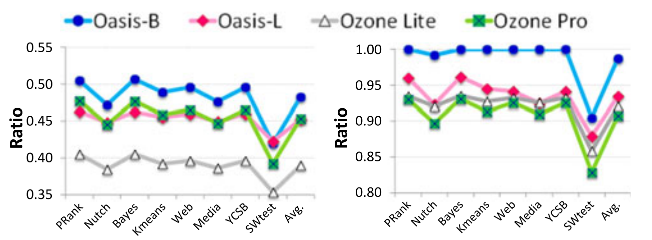
7.1 OpEx Improvement

In additional to electricity fee, the utility companies also charge a power demand fee at about \$8~9/kW per billing



(a) High RES Variability (b) Low RES Variability

Fig. 13. The average battery backup capacity. Ozone maintains high backup time due to its better battery capacity management capability.



(a) High RES Variability (b) Low RES Variability

Fig. 14. The ratio of green energy usage to overall power consumption. Ozone Pro has relatively higher dependence on grid power.

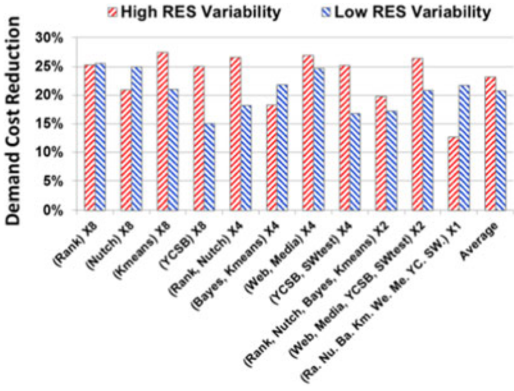


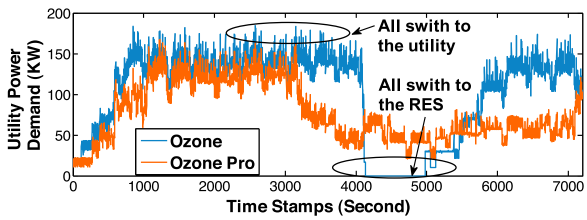
Fig. 15. The peak power cost savings of Ozone Pro.

cycle [17], [18]. Fig. 15 presents the monthly power demand cost savings for various workload combinations under different renewable power generations. We consider both homogeneous *Oasis Nodes* that run the same workloads and heterogeneous *Oasis Nodes* that run different workloads. With *Ozone Pro*, one can reduce power demand cost by 23 percent for High RES Variability.

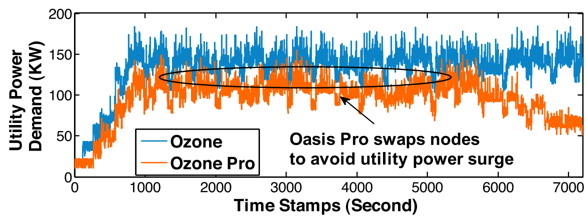
The driven insight behind *Ozone*'s global coordination is that the switching of *Oasis Nodes* should be carefully planned to avoid utility power surge. As shown in Fig. 16, it is likely that large amounts of *Oasis Nodes* simultaneously enter into battery charging state, causing high utility power demand. Once these nodes are fully charged, they may again simultaneously switch back to the solar power side, causing abrupt decrease in utility power demand. This kind of operation is not cost-efficient and may cause scheduling difficulty for local utility company. Whether the variability of renewable power is high or not does not affect the optimization effectiveness of *Ozone Pro*. The cost reduction depends on many factors such as solar power output, load power demand, and initial battery capacity.

7.2 CapEx Improvement

We evaluate the cost of *Oasis* design (excluding the IT server cost and the cost of labor). Fig. 17 presents two pie charts that show the cost breakdown. The computing equipment is



(a) High RES Variability



(b) Low RES Variability

Fig. 16. Utility power traces under different scheduling schemes.

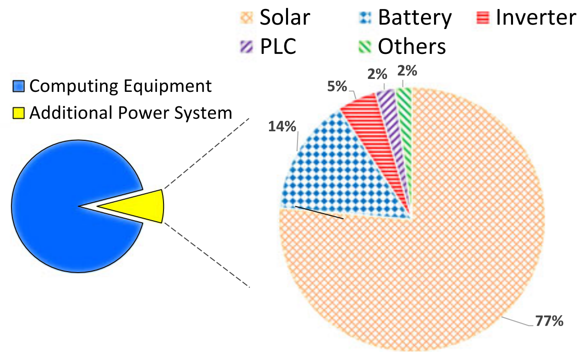


Fig. 17. Cost breakdown of the Oasis node scaled to a 40U standard rack of moderate power density (<10 KW).

the dominant component at the scale of a 40U standard rack. The additional power system accounts for about 8 percent of the overall *Oasis Node*. In specific, solar panel is the most expensive component (77 percent of the power system), followed by battery (14 percent), power inverter (5 percent), and PLC module (2 percent). The PLC and HMI systems account for only a small portion of the TCO since they do not need to scale up when more servers are added.

Centralized renewable power integration has relatively low initial cost due to the scale effect. Recent report estimates that small-scale PV panel (around 5 KW, for residential use) has an installed price of \$5.9/W, while large-scale PV panel (several hundred KW) has a lower price of around \$4.74/W [18]. In addition, solar power inverter accounts for about 10~20 percent of the initial system cost [19]. Central inverters in several MWs level are often cheaper compared to micro inverters (typically < 10 KW) used in the *Oasis* controller. The former costs around \$0.18/W, while the later costs around \$0.5/W [20].

The main advantage of *Oasis* is that it allows users to gradually increase the installed renewable power capacity. *Oasis* users can also take advantage of the ever-decreasing component cost to further lower their total expenditures. It has been shown that the installed prices of U.S. residential and commercial PV systems declined 5~7 percent per year, on average, from 1998~2011, and by 11~14 percent from 2010~2011, depending on system size [20]. The cost of micro-inverter also decreases by 5 percent yearly [20]. In addition, prices for energy storage system used to support renewables integration will fall in coming years. The total initial cost for lead acid batteries is \$784~\$2383 per kWh, with a yearly declining rate of 5~12 percent. The total initial cost for Lithium batteries is \$715~\$1638 per kWh, with a yearly declining rate of 10~14 percent [21].

Fig. 18 illustrates how *Oasis* design helps to improve the overall cost-effectiveness of renewable energy powered scale-out datacenters. It includes the initial setup cost of the green data center. We assume that *Oasis* users evenly increase their deployment of *Oasis Node* with a ten-year scale-out plan. (e.g., equip 10 percent of the datacenter servers with solar power system every year). For the cost of solar power system, we use a conservative decline rate of 6 percent per year, and an optimistic decline rate of 12 percent per year. We calculate electricity cost savings (\$0.1 per kWh) based on historical solar traces. We use hourly solar irradiance data (Jan 2003 ~ Dec 2012, 24 hours a day)

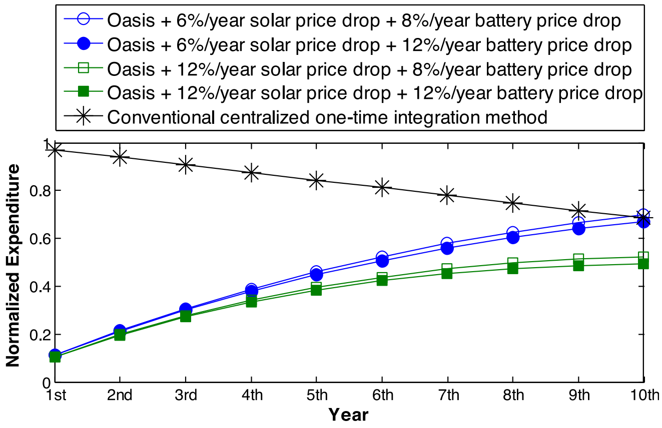


Fig. 18. The hardware and energy cost of deploying Oasis in a green scale-out datacenter.

provided by the NREL Solar Radiation Research Laboratory (http://www.nrel.gov/midc/srrl_bms/). We assume the utility power is \$0.1/kWh and datacenters can sell excess renewable power to the utility.

Fig. 18 shows the cost overhead (total additional cost due to green power integration) for different designs. The results are normalized to the one-time cost of datacenter-level solar power integration. For conventional centralized integration, one can expect to get 31 percent investment return (due to electricity savings) after 10 years. However, this estimation is optimistic as the utility grid typically uses negotiated renewable power feed-in tariff that has a lower purchase price. Although batteries do affect the TCO, the cost savings of *Oasis* is dominated by solar power systems. If solar panel cost decreases by a conservative rate of 6 percent per year, the total cost savings is close to a centralized design. If solar power cost declines faster (e.g., 12 percent per year), *Oasis* could result in 26 percent less overhead cost.

There is no one-for-all universal design. Green datacenters with centralized solar system may be a better choice if one has a firm goal of renewable power integration capacity and a confidence in future load demand. However, if green datacenter operators want flexibility and seek to gradually expand to avoid over-committing the capital, *Oasis* provides an attractive alternative.

8 Related Work

The scale-out model starts to draw considerable attention in datacenters. At the processor level, there has been several papers that introduce design methodologies for scale-out systems [22], [23]. At the datacenter level, traditional approaches mainly emphasize improving server efficiency and density to gain additional power capacity [7], [24]. Differently, this work looks at incrementally offering additional green power to servers to enable them to scale out.

Prior work on green energy aware systems can be classified into either simulation-based or prototype-based:

8.1 Simulation-Based Design

We observe three phases in the design and management of green energy powered systems. At first, designers mainly focus on hardware and system control. For example, one can leverage server power scaling to track the time-varying renewable power budget [5]. By leveraging grid-tie inverters, one can also achieve intelligent green energy routing [25]. The second stage features more flexible solutions that leverage various load adaptation schemes [26]. The main idea is to shift deferrable jobs to a time window in which renewable power generation is sufficient, or to relocate workloads to a different system where power budget is abundant [27]. In the third stage, the gap between energy source management and workload management starts to diminish. One can cooperatively tune energy sources and workloads to achieve the desired tradeoffs [4], [28], [29].

Another representative group of related work is in the context of resource management. For example, by intelligently control the workload, one can reduce network energy demand [30] or minimize the peak inlet temperature to lower cooling power needs [31]. Prior work has shown that leveraging renewable energy usage can lower datacenter costs [32], [33], [34]. Integrating nonconventional storge systems (thermal storage, supercapacitors) managing workload accordingly can further improve datacenter sustainability [35], [36], [37].

Prior proposals typically assume that the interface between energy source and server system is ready. Although future smart grid is expected to feature smart gateway for providing connectivity and interactive control between onsite power generator and computing load, currently such interface is not widely adopted.

8.2 Prototype-Based Design

Several studies have demonstrated the feasibility of renewable energy powered green datacenters. In Table 4 we compare *Oasis* with three representative systems. Blink [38] is a power management scheme for server clusters running on

TABLE 5
Comparison of Representative Green Datacenter Prototypes

Catagory \ System	Blink	Net-Zero	Parasol	Oasis
Power Source	Solar + wind	Grid + Solar	Grid + Solar	Grid + Solar
RES Integration	Not synchronized	Synchronized	Synchronized	Not synchronized
Battery System	Centrally installed	Not installed	Centrally installed	Highly distributed
Workload Type	Interactive	Critical/Non-critical	Deferrable/Non-deferrable	Load-independent
Opt. Objective	Near-optimal hit rate	Net-zero energy usage	Cost and performance	Scalability and cost
Tuning Knobs	Server power state	Job execution schedule	Grid-tie, Pwr. states	Switch, Perf. states
Granularity	Per cluster	Per cluster	Per container	Per Oasis node

TABLE 6
Comparison of the First and Second Generation Design of OASIS

Catagory \ System	*Oasis* 1st Generation (1G)	*Oasis* 2nd Generation (2G)
Management Framework	*Ozone Lite* running on external node	*Ozone Pro* based on the OpenStack platform
Automated Configuration	Not Supported	Automated configuring, metering, and logging
Multi-node Management	Not Supported	Global power switching coordination
Energy Buffer Control	Strict energy throughput capping	Flexible energy throughput capping

standalone renewable energy systems. It can intelligently tune the on/off cycles of server motherboard to track wind power output. Net-Zero [39] is a renewable energy powered micro-cluster designed by HP for minimizing datacenter's dependence on traditional utility grid. Its scheduling algorithm focuses on managing non-critical workload. The Parasol project involves a series of designs: GreenSolt [40], GreenHadoop [27], and GreenPar [41]. Its prototype is a solar-powered micro-data center backed by utility grid and batteries. It also highlights free-air cooling, Atom-based servers, renewable power prediction, etc.

Our work differs from prior studies in many aspects. For example, *Oasis* focuses on how to incrementally expand the capacity of a green datacenter and how to power-manage newly added components. Differently, prior designs mainly targets datacenter-level renewable energy integration and control for warehouse-scale or container-scale systems [42]. In addition, *Oasis* does not synchronize solar power with the utility power. It relies on highly distributed battery systems to integrate renewable energy. This allows for both fine-grained renewable power expansion and heavy renewable energy penetration. Moreover, *Oasis* introduces ways to optimize system operation through a joint power supply switching and performance state tuning. Prior works mainly emphasize temporal/spatial job scheduling and server power state tuning.

We have extended the functionality of *Oasis* significantly compared to its initial prototype [43]. Table 5 summarizes the key differences between the first and second model of *Oasis*. *Oasis* 1G is managed by the Lite version of *Ozone*. It only supports autonomous switching mode and very basic battery management policies running on an external node. In contrast, *Oasis* 2G uses a more professional version of *Ozone*. Its OpenStack-based central management platform enables automated configuration, multi-node coordination, and more flexible battery control. As a result, *Oasis* 2G could provide better system expansion capability, lower TCO, and higher average workload performance.

9 Conclusion

In this paper we envision the long-term competitiveness of introducing unconvential power provisioning architecture into datacenter capacity expansion plan. We show that integrating green energy powered modular server racks allows a datacenter to scale out with certain specific goals, such as diversification of power sources, carbon emission capping, and saving cost. With smart global and local power management on renewable power supplies and distributed energy storage systems, one can achieve better design tradeoff in a green datacenter.

Acknowledgments

This work is supported in part by NSF 1320100, 1117261, 0937869, 0916384, 0845721(CAREER), 0834288, 0811611, 0720476, by NSFC grant 61128004 and 61502302. Chao Li is also supported by a CCF-Intel Young Faculty Researcher Program Award and a CCF-Tencent Open Fund.

References

[1] C. Belady, "Projecting annual new datacenter construction market size," White Paper, Microsoft, http://docplayer.net/7092037-Projecting-annual-new-datacenter-construction-market-size-by-christian-l-belady-p-e.html, 2011.

[2] Digital Realty Trust, "Research study indicates widespread datacentre expansions by European companies during next 24 months," http://www.datacentres.com/news/research-study-indicates-widespread-datacentre-expansions-european-companies-during-next-24, 2007.

[3] The Uptime Institute, "The 2012 uptime institute data center industry survey," White Paper, http://www.egsa.org/Portals/7/Documents/CommitteeDocs/Misc/Uptime_Institute_2012_Survey.pdf, 2012.

[4] Í. Goiri, W. Katsak, K. Le, T. D. Nguyen, and R. Bianchini, "Parasol and greenswitch: Managing datacenters powered by renewable energy," *ACM SIGARCH Comput. Archit. News*, vol. 41, no. 1, pp. 51–64, 2013.

[5] C. Li, W. Zhang, C. Cho, and T. Li, "SolarCore: Solar energy driven multi-core architecture power management," in *Proc. IEEE 17th Int. Symp. High Perform. Comput. Archit.*, 2011, pp. 205–216.

[6] H. Sher and K. Addoweesh, "Micro-inverters—Promising solutions in solar photovoltaics," *Energy Sustain. Develop.*, vol. 16, no. 4. pp 389–400, 2012.

[7] V. Kontorinis, et al., "Managing distributed UPS energy for effective power capping in data centers," in *Proc. 39th Annu. Int. Symp. Comput. Archit.*, 2012, pp. 488–499.

[8] P. Sarti, "Battery cabinet hardware v1.0," White paper, Open Compute Project, 2012.

[9] M. Ton, B. Fortenbery, and W. Tschudi, "DC power for improved data center efficiency," Technical Report of Lawrence Berkeley National Laboratory, 2007.

[10] N. Jaalam, N. Rahim, A. Bakar, C. Tan, and A. Haidar, "A comprehensive review of synchronization methods for grid-connected converters of renewable energy source" *Renewable Sustain. Energy Rev.*, vol. 59, pp. 1471–1481, 2016.

[11] Ponemon Institute, "National survey on data center outages," White Paper, http://www.emersonnetworkpower.com/documentation/en-us/brands/liebert/documents/white%20papers/data-center-optimization_24655-r09-10.pdf, 2010.

[12] H. Bindner, T. Cronin, P. Lundsager, J. Manwell, U. Abdulwahid, and L. Baring-Gould, "Lifetime modelling of lead acid batteries," *Risø Nat. Laboratory*, Roskilde, Denmark, Tech. Rep. R-1515, 2005.

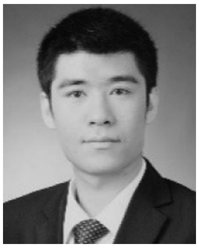
[13] D. Doerffel and S. Sharkh, "A critical review of using the Peukert equation for determining the remaining capacity of lead-acid and lithium-ion batteries," *J. Power Sources*, vol. 155, no. 2, pp. 395–400, 2006.

[14] S. Huang, J. Huang, J. Dai, T. Xie, and B. Huang, "The HiBench benchmark suite: Characterization of the MapReduce-based data analysis," *New Frontiers Inf. Softw. Serv.*, vol. 74. pp. 209–228, 2011

[15] M. Ferdman, et al., "Clearing the clouds: A study of emerging scale-out workloads on modern hardware," in *Proc. 17th Int. Conf. Archit. Support Program. Languages Operating Syst.*, 2012, pp. 37–48.

[16] Understanding Electric Demand, White Paper, Nat. Grid, 2012. [Online]. Available: https://www.nationalgridus.com/niagaramohawk/non_html/eff_elec-demand.pdf
[17] Residential Demand Rate. (2016). [Online]. Available: http://www.georgiapower.com/pricing/residential/residential-demand.cshtml
[18] D. Feldman, G. Barbose, R. Margolis, R. Wiser, N. Darghouth, and A. Goodrich, "Photovoltaic (PV) pricing trends: Historical, recent, and near-term projections," Nat. Renewable Energy Laboratory, Golden, CO, USA, Tech. Rep. DOE/GO-102012-3839, 2012.
[19] "A review of PV inverter technology cost and performance projections" Nat. Renewable Energy Lab, Golden, CO, USA, Tech. Rep. NREL/SR-620-38771, 2006.
[20] R. Simpson, "Levelized cost of energy from residential to large scale PV," http://www.psma.com/sites/default/files/uploads/tech-forums-alternative-energy/presentations/2012-apec-256-evolution-photovoltaic-inverter.pdf, 2012.
[21] LAZARD, "LAZARD's levelized cost of storage analysis—version 1.0," https://www.lazard.com/media/2391/lazards-levelized-cost-of-storage-analysis-10.pdf, 2015.
[22] P. Lotfi-Kamran, et al., "Scale-out processors," in *Proc. 39th Int. Symp. Comput. Archit.*, 2012, pp. 500–511.
[23] S. Li, K. Lim, P. Faraboschi, J. Chang, P. Ranganathan, and N. Jouppi, "System-level Integrated Server Architectures for Scale-out Datacenters," in *Proc. 44th Int. Symp. Microarchitecture*, 2011, pp. 260–271.
[24] D. Wang, C. Ren, and A. Sivasubramaniam, "Virtualizing power distribution in datacenters," in *Proc. 40th Int. Symp. Comput. Archit.*, 2013, pp. 595–606.
[25] N. Deng, C. Stewart, and J. Li, "Concentrating renewable energy in grid-tied datacenters," in *Proc. Int. Symp. Sustain. Syst. Technol.*, 2011, pp. 1–6.
[26] Í. Goiri, et al., "Leveraging green energy in data-processing frameworks," in *Proc. 7th ACM Eur. Conf. Comput. Syst.*, 2012, pp. 57–70.
[27] C. Li, A. Qouneh, and T. Li, "iSwitch: Coordinating and optimizing renewable energy powered server clusters," in *Proc. 39th Int. Symp. Comput. Archit.*, 2012, pp. 512–523.
[28] C. Li, R. Zhou, and T. Li, "Enabling distributed generation powered sustainable high-performance data center," in *Proc. 19th Int. Symp. High Performance Comput. Archit.*, 2013, pp. 35–46.
[29] C. Li, R. Wang, D. Qian, and T. Li, "Managing server clusters on renewable energy mix," *ACM Trans. Auton. Adaptive Syst.*, vol. 11, no. 1, 2016, Art. no. 1.
[30] D. Li, Y. Shang, W. He, and C. Chen, "EXR: Greening data center network with software defined exclusive routing," *IEEE Trans. Comput.*, vol. 64, no. 9, pp. 2534–2544, 2015.
[31] Q. Tang, S. Gupta, and G. Varsamopoulos, "Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach," *IEEE Trans. Parallel Distrib. Syst.*, vol. 19, no. 11, pp. 1458–1472, 2008.
[32] C. Ren, D. Wang, B. Urgaonkar, and A. Sivasubramaniam, "Carbon-aware energy capacity planning for datacenters," *Proc. 20th Int Symp. Modelling Anal. Simul. Comput. Telecommun. Syst.*, 2012, pp. 391–400.
[33] F. Kong, X. Liu, and L. Rao, "Optimal energy source selection and capacity planning for green datacenters," in *Proc. ACM Int. Conf. Meas. Model. Comput. Syst.*, 2014, pp. 575–576.
[34] 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, 2015.
[35] 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.*, 2014, vol. 25, no. 8, pp. 2030–2042.
[36] L. Liu, et al., "Leveraging heterogeneous power for improving datacenter efficiency and resiliency," *IEEE Comput. Archit. Lett.*, vol. 14, no. 1, pp. 41–45, Jan.-Jun. 2015.
[37] L. Liu, et al., "HEB: Deploying and managing hybrid energy buffers for improving datacenter efficiency and economy," *ACM SIGARCH Comput. Archit. News*, vol. 43, no. 3, pp. 463–475, 2015.
[38] N. Sharma, S. Barker, D. Irwin, and P. Shenoy, "Blink: Managing server clusters on intermittent power," *ACM SIGPLAN Notices*, vol. 47, no. 4, pp. 185–198, 2011.
[39] P. Banerjee, C. Patel, C. Bash, A. Shah, and M. Arlitt, "Towards a net-zero data center," *ACM J. Emerg. Technol. Comput. Syst.*, vol. 8, no. 4, 2012, Art. no. 27.
[40] Í. Goiri, et al., "Greenslot: Scheduling energy consumption in green datacenters," in *Proc. Int. Conf. High Performance Comput. Netw. Storage Anal.*, 2011, pp. 1–11.
[41] M. Haque, I. Goiri, R. Bianchini, and T. Nguyen, "GreenPar: Scheduling parallel high performance applications in green datacenters," in *Proc. 29th Int. Conf. Supercomputing*, 2015, pp. 217–227.
[42] C. Li, et al., "Towards sustainable in-situ server systems in the big data era," *ACM SIGARCH Comput. Archit. News*. vol. 43, no. 3, pp. 14–26, 2015.
[43] C. Li, et al., "Enabling datacenter servers to scale out economically and sustainably," in *Proc. 46th Annu. IEEE/ACM Int. Symp. Microarchitecture*, 2013, pp. 322–333.

Chao Li received the BS degree from Zhejiang University, in 2009 and the PhD degree from the University of Florida, in 2014. He is currently a Tenure-Track assistant professor in the Department of Computer Science and Engineering, Shanghai Jiao Tong University. His research interests include computer architecture, green data center, and emerging computing platforms. He has received the Best Paper Award from IEEE HPCA in 2011 and the *IEEE Computer Architecture Letter* in 2015. He is a member of the IEEE.



Yang Hu received the BS degree from Tianjin University, in 2007 and the MS degree from Tsinghua University, in 2011. He is currently working toward the PhD degree in IDEAL lab at the University of Florida. His research interests include data center, green cloud computing, and cloud computing.

Juncheng Gu received the BS degree from Central South University, China, in 2013 and the MS degree from the University of Florida, in 2015. He is currently working toward the PhD degree at the University of Michigan, Ann Arbor. His research interests include green computing and cloud computing.

Jingling Yuan received the BS, MS, and PhD degrees from Wuhan University of Technology, China. She was a visting scholar at the University of Florida, in 2009. She is current a professor in the School of Computer Science and Technology, Wuhan University of Technology. Her research interests include multi-core computing, green computing, big data, and machine learning.

Tao Li received the PhD degree from the University of Texas at Austin. He is a professor in the Department of Electrical and Computer Engineering, University of Florida. His research interests include computer architecture, the impacts of emerging technologies/applications on computing, and evaluation of computer systems. He received the 2009 NSF CAREER Award, the 2008, 2007, 2006 IBM Faculty Awards, the 2008 Microsoft Research Safe and Scalable Multicore Computing Award, and the 2006 Microsoft Research Trustworthy Computing Curriculum Award. He is general chair of HPCA 2014.