Enabling Datacenter Servers to Scale Out Economically and Sustainably

Chao Li*, Yang Hu*, Ruijin Zhou*, Ming Liu*, Longjun Liu§, Jingling Yuan†, and Tao Li*

* Intelligent Design of Efficient Architectures Laboratory (IDEAL)
Department of Electrical and Computer Engineering
University of Florida, USA
{chaol, huyang.ece, zhourj, mingliu}@ufl.edu, taoli@ece.ufl.edu

§ Xi'an Jiaotong University, China
longjun.liu@stu.xjtu.edu.cn
† Wuhan University of Technology, China
yjl@whut.edu.cn

ABSTRACT

As cloud applications proliferate and data-processing demands increase, server resources must grow to unleash the performance of emerging workloads that scale well with large number of compute nodes. Nevertheless, power has become a crucial bottleneck that restricts horizontal scaling (scale out) of server systems, especially in datacenters that employ power over-subscription. When a datacenter hits the maximum capacity of its power provisioning equipment, the owner has to either build another facility or upgrade existing utility power infrastructure – both approaches add huge capital expenditure, require significant construction lead time, and can further increase the owner's carbon footprint.

This paper proposes *Oasis*, a power provisioning scheme for enabling power-/carbon- constrained datacenter servers to scale out economically and sustainably. *Oasis* naturally supports incremental power capacity expansion with near-zero environmental impact as it takes advantages of modular renewable energy system and emerging distributed battery architecture. It allows scale-out datacenter to double its capacity using 100% green energy with up to 25% less overhead cost. This paper also describes our implementation of *Oasis* prototype and introduces our multi-source driven power management scheme *Ozone*. *Ozone* allows *Oasis* to identify the most suitable power supply control strategies and adjust server load cooperatively to maximize overall system efficiency and reliability. Our results show that *Ozone* could reduce the performance degradation of *Oasis* to 1%, extend *Oasis* battery lifetime by over 50%, and almost triple the average battery backup capacity which is crucial for mission-critical systems.

Categories and Subject Descriptors:

C.5.5 [**Computer System Implementation**] Servers

General Terms: Design; Management; Experimentation.

Keywords: datacenter, scalability, green energy, sustainability, power management, energy storage, cloud workload

MICRO-46, December 07-11, 2013, Davis, CA, USA.

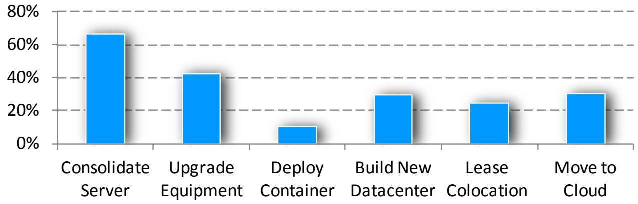


Figure 1: "How will your organization handle increased capacity demand over the next 12-18 months" – results from the Uptime Institute 2012 Data Center Industry Survey [3].

1. INTRODUCTION

The rapid adoption of cloud computing in today's economies is deemed a powerful engine for the growth of installed server capacity. To support emerging data-analytic workloads that tend to scale well with large number of compute nodes, modern datacenters are continually adding computing resources (i.e., scaling out) to their existing sites. In turn, global server market size is projected to triple in 2020, accounting for over 1000 TWh annual energy consumption [1, 2].

Over time, the constant influx of server resources in datacenters will eventually become *power-constrained.* According to a recent industry survey from the Uptime Institute [3], 30% of the enterprise datacenter managers expected to run out of power capacity within 12 months. Figure 1 shows several widely adopted solutions to the power capacity problem. While leasing collocated racks and deploying cloud services fit small server clusters on a budget, this study mainly focuses on large-scale enterprise datacenters and approaches that aim at improving computing capability and capacity. Server consolidation is mature techniques that can free up power capacity. However, in power-constrained datacenters, even consolidated servers have to limit their performance using either software-based (i.e., virtual CPU allocation) or hardware-based (i.e., DVFS) control knobs to avoid tripping circuit-breakers and causing costly downtime.

Upgrading power systems is a radical solution that allows one to add more servers, racks, and even onsite containers. However, like building a new datacenter, re-sizing datacenter initial power capacity can be a great undertaking since conventional centralized power provisioning scheme does not scale well. In a typical datacenter, the power delivery path involves multiple power equipments across several layers, as shown in Figure 2. Upgrading power infrastructure often requires re-design of the entire power delivery path which is not only costly but also time-consuming. Worse, utility power feeds are often at their capacity in some urban areas and additional electricity access for datacenter is restricted.

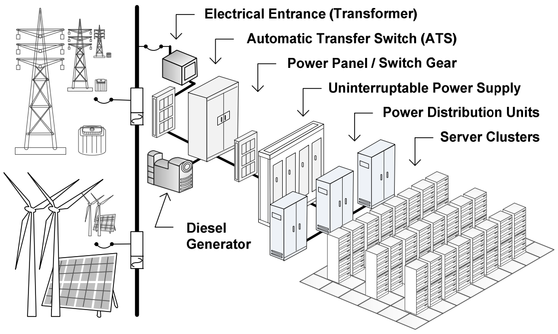


Figure 2: Conventional datacenter power architecture is not 'scale-out friendly'. Existing facility-level renewable power integration can also be costly and problematic.

In addition, modern scale-out datacenters are not only *power-constrained*, but also *carbon-constrained*. As server power demand increases, the associated carbon footprint expansion poses significant challenges for scale-out datacenters. It has been shown that global greenhouse gas (GHG) emissions could easily exceed 100 million metric tons per year if we keep using utility grid power that is generated from conventional fossil fuel [4]. An emerging solution to the carbon reduction problem is to leverage green energy sources to lower the environmental impact of computing systems. Several companies, including Microsoft, IBM, Google, Apple, and eBay, all start to explore renewable energy in datacenters as part of their long-term energy strategies and corporate environmental goals [5-9]. For example, eBay is experimenting with a small datacenter powered by a 100kW solar array. Apple's Maiden datacenter in North Carolina draws renewable power from both self-generation (60%) and regional plant (40%).

Unfortunately, to the best of our knowledge, existing green energy integration schemes often employ centralized power integration architecture and largely overlook the modularity of renewable power supplies. As shown in Figure 2, one of the advantages of such facility-level integration is that the renewable power supplies can be synchronized to the utility grid to eliminate the negative impact of renewable power variation on datacenter servers. However, the operation of such grid-connected renewable power system often relies on the availability of utility power. Meanwhile, the centralized power integration not only results in single-point of failure, but also makes future capacity expansion expensive.

In this study we ask one important question: *"Can we enable a power-constrained datacenter to gracefully scale out while lowering the carbon footprint with high efficiency and low overhead?"* Faced with the ever-growing computing demand, the skyrocketing server power expenditures, and the looming global environmental crisis, this question is of significant importance to datacenter operators who have to embrace efficient power provisioning and management schemes to survive economically and sustainably.

We propose *Oasis*, a unified power provisioning framework that synergistically integrates energy source control, power system monitoring, and architectural support for power-aware computing. *Oasis* leverages modular renewable power integration and distributed battery architecture to provide flexible power capacity increments. It allows power-/carbon- constrained systems to stay on track with horizontal scaling and computing capacity expansion.

We implement *Oasis* as a prototyped research platform for exploring key design tradeoffs in multi-source powered datacenter. Our first generation prototype is a micro server rack (12U) that draws power from onsite solar panels, conventional utility power, and local energy storage devices. These energy sources are coordinated by a micro-controller based power control hub (PCH) built from scratch. We customize the PCH to make it a rack-mounted, interactive system for easy installation and diagnosis.

We further propose *Ozone*, a power management scheme for *Optimized Oasis Operation* (O_3). Transcending the boundaries of traditional primary and secondary power, *Ozone* creates a smart power source switching mechanism that enables server system to deliver high performance while maintaining desired system efficiency and reliability. *Ozone* is able to dynamically distill crucial runtime statistics of different power sources to avoid unbalanced usage of different power sources. It could identify the most suitable control strategies and adaptively adjusts the server speed via dynamic frequency scaling to maximize the benefits of *Oasis* system.

To summarize, *Oasis* and *Ozone* constitute an interesting research platform from many perspectives: (1) it enables datacenter power system to scale-out and facilitates partial green power integration; (2) it links power supply management and server system control, and could enable real-time power supply driven workload scheduling; (3) its power provisioning architecture (i.e., hybrid power supplies + distributed control domain) could bring us improved datacenter availability in many power failure scenarios; (4) its decentralized multi-source power management could provide datacenter operator the flexibility of offering different green computing services based on different customer expectations.

This paper makes the following contributions:

- We propose *Oasis*, a novel power provisioning architecture that enables server power supply to scale-out, thereby facilitating initial capacity planning and on-demand datacenter capacity expansion.
- We build a prototype system of *Oasis* from scratch. The system provides interactive communication portal between hybrid power supply and server system, thereby enabling real power-driven workload control and workload-driven energy source management.
- We propose *Optimized Oasis Operation (Ozone)* that could jointly optimize battery service life, battery backup time, and workload performance. Our evaluation results show that *Oasis* is able to further reduce workload execution delay to 1%, extend battery lifetime by over 50%, increase battery autonomy time by 1.9X, while still maintains satisfactory green energy usage rate.
- We further discuss the impact of *Oasis* on the cost of large-scale datacenter design. We show that *Oasis* could enable scale-out datacenters to double its capacity with up to 29% less cost overhead, compared to the facility-level one-time renewable energy integration.

The rest of this paper is organized as follows. Section 2 illustrates *Oasis* topology and describes design rationale. Section 3 presents the implementation of *Oasis*. Section 4 proposes *Ozone* power management. Second 5 introduces experimental methodology. Section 6 presents our evaluation of *Ozone*. Section 7 analyzes the cost-effectiveness of *Oasis*. Section 8 discusses related work and Section 9 concludes this paper.

2. OASIS: OVERVIEW AND RATIONALE

In this section we first give an overview of typical datacenter expansion strategies, which we refer to as power capacity scale-out models. We then introduce modular power sources that could be leveraged to facilitate efficient and flexible capacity expansion. Finally, we discuss important design considerations for *Oasis*.

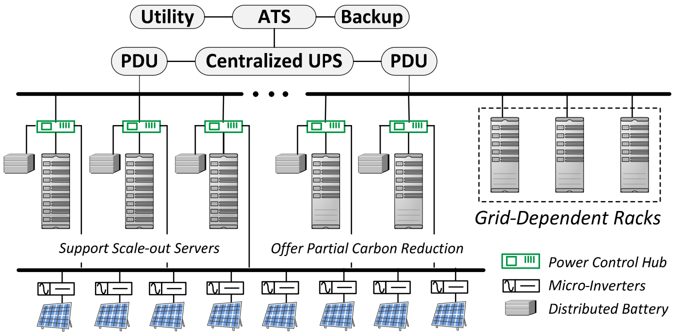


Figure 3: *Oasis* power provisioning architecture leverages the modularity of both battery systems and renewable power systems.

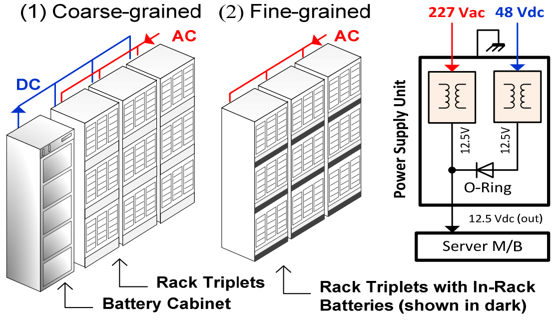


Figure 4: Distributed energy storages (battery) emerge in recent implementation of server clusters [11].

2.1 Scale-Out Models

We classify existing power capacity scale-out model as either *utility power over-provisioning* or *centralized power capacity expansion*. In the first case, the utility power and datacenter power delivery infrastructure is designed to support the maximal power the load may ever draw. Although such design provides abundant power headroom for scale-out server system, it inevitably increases the carbon footprint. In the second case, the power delivery infrastructure is provisioned for a certain level of anticipated future power drawn and the scaling out is handled entirely by datacenter-level power capacity integration. However, installing large-scale renewable power system often results in high expansion cost.

Oasis explores a "pay-as-you-grow" model for scale-out datacenters, which we refer to as *distributed green power increments* model. In this model, the utility grid and datacenter power delivery infrastructure are provisioned for a fixed level of load power demand. When datacenter reaches its maximum capacity, however, 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 its capitals. As shown in Figure 3, *Oasis* consists of a number of green energy powered computing racks, called *Oasis nodes*. Each *Oasis node* is attached to a distributed power control hub (PCH), which further connects to onsite renewable power supply, power distribution unit (PDU) for utility power, and local distributed energy storage devices. In Section 2.2 we introduce modular power sources that allow us to add power capacity incrementally; in Section 2.3 we discuss the rationale for distributed green power integration in datacenters.

2.2 Modular Power Sources

Scale-out datacenters need both modular standby power sources that provide incremental backup power capacity and modular primary power sources that generate additional electrical power. Today, distributed battery architecture is emerging to improve datacenter efficiency [10, 11]. It provides us a convenient way to add additional backup power as computing demand increases. In addition, renewable power supplies such as solar panels are usually modular and highly scalable in their capacity. They are ideal power supplies that support carbon-free server scale-out.

2.2.1 Distributed Energy Storage System

Google and Facebook propose employing distributed energy storage to avoid the energy efficiency loss due to power double-conversion in conventional centralized UPS systems [10, 11]. Such de-centralized design can also avoid single-point of failure and increase overall datacenter availability. Recently, the distributed battery topology has been further used to shave peak power to free up datacenter capacity [10]. In this study, we leverage distributed battery for supporting incremental server system scale-out.

Figure 4 illustrates the distributed battery provisioning architectures in the Open Compute Project [11] led by Facebook. In the coarse-grained integration scheme, a battery cabinet populated with commodity lead-acid batteries is used to provide standby power for a rack triplet. Each triplet consists of three column racks and each rack is further divided into three power zones. The battery cabinet also includes several breakers, quick fuse disconnects, sensors, and a high current DC bus bar. In the fine-grained integration scheme, the battery cabinet is replaced by a high-density lithium-ion battery backup unit (BBU) in each rack power zone. In both cases, battery provides 48V DC backup power to the servers and could provide around 45 seconds of runtime at full load [11].

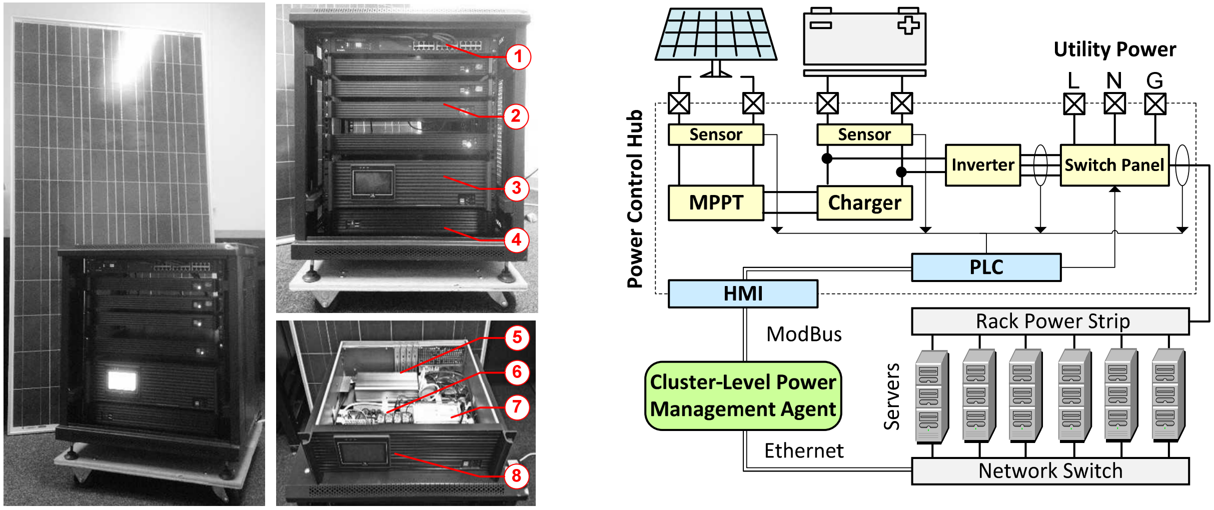
2.2.2 Solar Power with Micro-Inverters

Wind turbine and solar panel are both modular power sources. Compared to wind turbines, solar panels can provide even smaller capacity increments. In this study we look at solar panels that use micro-inverter [12] to provide incremental power capacity.

Conventionally, solar power system uses string inverters which require several panels to be linked in series to feed one inverter. String inverters are big, prone to failure from heat, and shows low efficiency. In contrast, micro-inverters are smaller in size and can be built as an integrated part of the panel itself. As shown in Figure 3, each panel has its own micro-inverter; solar panels are further connected in parallel to offer larger power capacity. Compared to centralized design, solar panel with micro-inverter shows much better scalability, reliability and efficiency. The only disadvantage is its relatively high cost per watt compared to centralized inverter. However, the amortized cost can be lower than a centralized system (detailed in Section 7).

2.3 Distributed Integration

Oasis adds renewable power capacity at power distribution unit (PDU) level. As shown in Figure 3, the PUD-level renewable energy integration 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 PDU, 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.



(a) System prototype (b) Node architecture

Figure 5: Prototype of *Oasis* node. Several key components include: (1) network switch, (2) computing nodes, (3) power control hub, (4) battery chassis, (5) converter and inverter, (6) power switch, (7) PLC, and (8) HMI.

Oasis does not synchronize renewable power to utility power due to three considerations. First, at the PDU level, renewable power synchronization induces voltage transients, frequency distortions, and harmonics. Datacenter servers are susceptible to the power anomalies. As the impact of power quality issue on server efficiency and reliability is still an open question, massively adding grid-tied renewable power at rack level is debatable. Second, even if one chooses to synchronize renewable power at datacenter facility-level, grid-tied renewable power system can cause efficiency problems. The newly added renewable power incurs many levels of redundant power conversion before reaching server racks, resulting over 10% energy loss [13]. Third, as required by UL 1741 and IEEE 1547, all grid-tied inverters must disconnect from the grid if they detect power islanding. That is, these renewable power systems must shut down if the grid power no longer present. As reported in a recent survey, the U.S. datacenter experiences 3.5 times of utility power loss per year with an average duration over 1.5 hours [14]. Thus, with synchronization based power provisioning, datacenters may lose renewable power supply even when they need it most.

The intention of *Oasis* project is to facilitate server scale-out in power-/carbon- constrained datacenters. *Oasis* aims to become a non-intrusive, modularized renewable power integration scheme which allows datacenter operators to increase power capacity on-demand and reduce datacenter carbon footprint gradually. In Section 3 we describe our implementation of *Oasis* and in Section 4 we propose smart power management scheme *Ozone* which allows *Optimized Oasis Operation* (O_3).

3. IMPLEMENTATION

In this section we describe our implementation of *Oasis* starting from scratch (Figure 5-a). Specifically, we focus on *Oasis nodes*, the key element of *Oasis* design. As shown in Figure 5-b, its structure consists of a power control hub, a server rack, and a power management agent that coordinates them.

3.1 Power Control Hub

The power control hub (PCH) is the major hardware addition of *Oasis nodes*. It integrates battery charger, inverter, power supply switch panel, programmable logic controller (PLC), and a human-machine interface (HMI) that allows easy system diagnosis. The PCH is designed to manage multiple energy sources (i.e., renewable power, utility power, and distributed battery devices). On its backside we provide three electrical sockets that allows easy connection to utility power (AC), battery (DC), and solar panels (DC).

Our prototype system is powered by an array of solar module. Each module is a 270 Watt Polycrystalline panel from Grape Solar. The panel output is a complex function of the solar irradiation, ambient temperature, and the load connected to it. To harvest the solar energy, we use a maximal power point tracker (MPPT) in our power control hub. The MPPT samples the output of the solar cells and applies proper control to obtain the optimal solar energy generation. However, it is not an essential part of the PCH. The MPPT is often a build-in module of the solar micro-inverter.

Oasis leverages distributed battery devices to provide temporary solar energy storage. Energy storage devices, such as uninterruptable power supplies (UPS), are widely used in datacenters to address the risk of interruptions to the main grid. We have customized a stack of nine 2Ah sealed lead-acid batteries for the power control hub (scalable in the future). Our battery chassis can provide 5~10 minutes of backup time, depending on the server load.

We convert the DC solar power to AC to match the output level of utility power. The AC solar power and AC utility power are merged (but isolated) at a switch panel in 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 which further feeds a cluster of server nodes. The entire rack is either powered by the utility power or renewable power, depending on the status of the internal power switch.

Table 1 shows several key technical data of our current *Oasis* prototype. The maximum charging current of our system is 4A, which is limited by the battery charger for battery lifetime considerations. The PCH itself consumes static power (about 9 Watts) due to the HMI and PLC operations. The dynamic power loss is due to the heat dissipation of power conversion systems. The lifetime of battery system varies depend on its charging/discharging profile. We have designed power management scheme to maximize the benefits of battery and optimize its service life (detailed in Section 4).

Maximum charging current	4 A
Static power loss	9 W
Dynamic power loss	8.8 W
Battery lifetime (Estimated)	3 ~10 Years

Table 1: The measured *Oasis* technical data

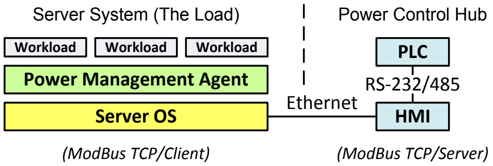


Figure 6: Power management agent and the data communication scheme inside of *Oasis node*

3.2 System Monitoring and HMI

We enrich *Oasis* with sensing components that keep monitoring real-time solar power output voltage and battery terminal voltage. These hardware agents could inform us the renewable energy resource conditions and the health status of batteries. Our systematic checkup of power supply behaviors also offers a real-time profile of the system's energy utilization. This feature makes it possible to pinpoint areas of high energy usage in server farms, while establishing a baseline for further capacity planning and optimization. Given appropriate monitoring data, *Oasis* could enable:

1) *Workload application driven energy source switching*
2) *Energy source driven server load adaptation*
3) *Emergency handling when facing power anomalies*

To facilitate real-time configuration and diagnosis, we have designed human-machine interface (HMI) for each power control hub. As shown in Figure 5-(a), it is a touch screen panel with build-in microprocessors that could display graphics, animation, and interchange data to and from the PLC. Besides, the HMI device also serves as an important portal for communication between the PLC and our external power management agent.

3.3 Bridging Server and Power Supply

A key feature of the PCH is that it sets up the communication gateway between power supply layer and server workload layer. This design is partly enlightened by the energy management strategy in future smart grid, which focuses on intelligent control, communication, and coordination across electrical loads, power electric interfaces, and power generators. Figure 6 illustrates the data communication scheme of *Oasis*. The power management agent is a middleware lies between operating system and workloads.

We use Modbus protocol [15] for communication between the power management agent and the power control hub. It is a widely used serial communication protocol for industrial electronic devices due to its robustness and simplicity. There are several other benefits of using Modbus. First, the Modbus TCP protocol integrates Modbus instruction sets into existing TCP/IP protocol. Therefore we can take advantage of the scalability of Ethernet to easily scale up our deployment of nodes and share information among them. Second, there is no need to worry about the transmission failure of the Modbus instruction as the lower layer TCP/IP protocol has provided the redundant checksum. Typical Modbus TCP communication includes a Modbus server and a Modbus client. In this study, the power management agent is the Modbus client (master) which initiates the communication requests periodically through socket to the Modbus server (slave), i.e., the HMI.

3.4 Dynamic Energy Source Switching

Oasis nodes are able to dynamically switch between green power supply and utility power. The PCH offers two power switching modes: the autonomous mode and the coordinated mode.

Autonomous Mode: This is the default mode for *Oasis*. In this mode, *Oasis nodes* intend to run autonomously, i.e., switching the load between solar power and utility grid based on the given solar and utility power budget. The PLC in our power control hub defines two atomic modules: *SupplySense* and *SupplySwitch*, as shown in Figure 7. While the former module focusing on setting parameters that are used for making energy source management decision, the latter one executes energy source switch.

Coordinated Mode: *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.

3.5 Enhancing Power Switching Reliability

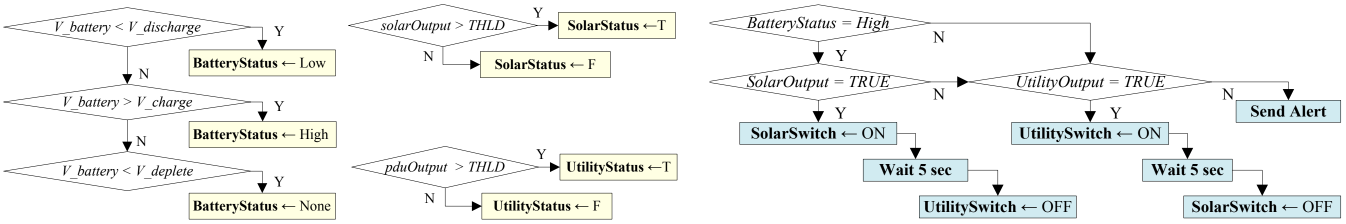
For every switch operation, the controller will first check the output of power supplies to ensure that they work normally. As shown in Figure 7, the PCH configuration allows an overlap between solar power and utility power when performing power supply switching. We disconnect one energy source only when the other has been successfully working for 5 seconds. This mechanism helps to avoid potential switching failure that interrupts server operation.

Our controller also maintains appropriate voltage thresholds to prolong the lifetime of backup power system. Figure 8 shows the measured battery voltage during power switching. The battery starts to discharge at 12.5V (charging threshold) and stop discharging when its voltage drops to 11.8V (discharging threshold). The charging/discharging thresholds prevent batteries from entering deep-discharging or over-charging mode. In addition, we notice that the large inrush current due to power switching could result in an immediate voltage droop (about 0.4V) in our battery pack, as shown in Figure 8. We are planning to design appropriate battery management circuitry to mitigate such abnormal battery drain issues. However, it is out of the scope of the themes pursued in this paper.

3.6 Server Power Demand Control

It is crucial to control the server power demand (especially the peak power drawn) when renewable energy generation fluctuates. For example, when solar power output decreases, one can temporarily lower the average processing speed of the server cluster to avoid shutting down any one of them. Although batteries can provide backup power capacity, it is beneficial to not heavily rely on them for reliability considerations.

In this study we use dynamic voltage and frequency scaling (DVFS) for server load adaption. It has been shown that DVFS could reduce the peak power drawn of a server cluster by 18% for real-world datacenter workload mix [16]. Our system kernel is configured with the on-demand frequency scaling governor. *Oasis* is able to set server system at different CPU operating states (C states) and performance states (P states). Based on the supply monitoring data and workload power demand statistics, *Oasis* performs real-time supply-driven load adaptation



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

Figure 7: The control flow of power supply switching. *Oasis* in autonomous mode uses two atomic modules to control the system.

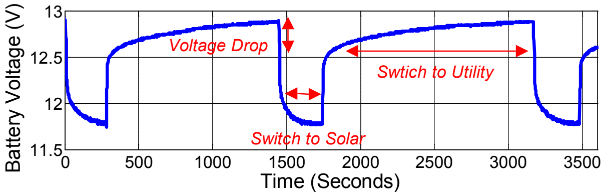


Figure 8: The power control hub uses threshold to avoid over charge/discharge a battery. The voltage droop is caused by inrush currents.

4. OPTIMIZED OASIS OPERATION

The flexibility of *Oasis* power provisioning architecture provides datacenter owners plenty of room for improving the efficiency and performance of their servers. In this section we propose *Ozone*, a power management scheme for *Optimized Oasis Operation* (O_3). *Ozone* enables *Oasis* users to take one step further towards the aim of efficient scale-out datacenters.

Ozone seeks a balance between power supply management and server load management. In addition to the basic power control rules of *Oasis*, *Ozone* features a new set of policies that intend to maximally take advantage of multiple energy sources without heavily relying on any single energy sources. In addition, *Ozone* adaptively adjusts server load based on power supply states to find the best design tradeoff.

While many factors may affect *Oasis* operation, *Ozone* picks up six representative factors for making control decisions. Three of them are power-related: *Utility Budget*, *Solar Budget*, *and Load Demand*; two of them are battery-related: *Discharge Budget* and *Remaining Capacity*; and one *Switch* parameter that specifies which energy source is in use as primary power.

4.1 Managing Battery Usage

Batteries have a lifespan (typically 5~10 years) estimated by their manufactures. These energy storage devices become no longer suitable for mission-critical datacenters after their designed service life. Aggressively discharging batteries will significantly degrade their lifetime. However, even if batteries are not frequently used, they still suffer aging and self-discharging problems.

Ozone uses an Ah-Throughput Model [17] for battery lifespan optimization. This model assumes that there is a fixed amount of aggregated energy (overall throughput) that can be cycled through a battery before it requires replacement. The model provides a reasonable estimation of the battery lifetime and has been used in software developed by the National Renewable Energy Lab [18].

During runtime, *Ozone* dynamically monitors battery discharging events and calculates battery throughput (in amp-hour) based on Peukets's equation [19]:

$$Discharge = I_{\text{actual}} \cdot (I_{\text{actual}} / I_{\text{nominal}})^{pc-1} \cdot T \qquad \text{(Eq-1)}$$

In Eq-1, I_{actual} is the observed discharging current, I_{norminal} is the nominal discharging current which is given by the manufacturer, pc is the Peukets coefficient, and T is the discharging duration. Over time, the aggregated battery throughput is given by:

$$D_{agg} = \sum_i Discharge_i \qquad \text{(Eq-2)}$$

To avoid battery over-utilization, *Ozone* carefully sets a soft limit on battery usage at any given time. At the beginning of each control cycle, *Oasis nodes* receive a *Discharge Budget* which specifies the maximum amount of stored energy uses that will not compromise battery lifetime. Assuming the overall battery throughput is D_{total}, the *Discharge Budget* is set as:

$$D_B = D_{\text{total}} - D_{\text{agg}} \qquad \text{(Eq-3)}$$

The *Discharge Budget* affects both power supply switching behavior and server workload adaptation control. When the *Discharge Budget* is inadequate, *Ozone* prefers to switch server back to utility power or decrease server power demand to avoid heavily utilizes battery system.

4.2 Managing Backup Capacity

Distributed battery systems not only provide *Oasis* basic support for managing time-varying renewable power, but also serve as uninterruptible power supplies for scale-out servers.

Maintaining necessary backup capacity is critical. It has been shown that UPS capacity exceeded is one of the top root causes of unplanned outages in datacenters [14]. The backup capacity is the primary factor that determines UPS autonomy (a.k.a. backup time). It is a measure of the time for which the UPS system will support the critical load during power failure. In addition to the *Discharge Budget*, *Ozone* also sets a limit on the minimum remaining capacity of batteries. Our system only uses 40% of the installed capacity for managing renewable power shortfall (referred to as *Flexible Capacity*). The remaining 60% battery capacity (i.e., *Reserved Capacity*) is only used for emergency handling purpose. When the battery capacity drops below 60%, *Ozone* will switch the server from renewable power supply to utility power supply.

4.3 Managing Server Performance

Due to the time-varying nature of renewable energy resources, *Oasis nodes* need to handle variable renewable power budget. There are always two power management options available: (1) decrease server performance level to lower server power demand, or (2) continue to operate at the highest frequency and use energy storage to compensate the power shortfall. To find the best design tradeoff between server performance and system reliability, *Ozone* cooperatively controls power supply switching and server processing speed, as shown in Table 2. The idea is to adaptively select one of the two power management options based on the observed *Discharge Budget* and the amount of *Flexible Capacity*.

In Table 2, when both *Discharge Budget* and *Flexible Capacity* are adequate, *Ozone* gives high priority to server performance boost (i.e., run workload at the highest frequency) with the support of battery. When the system runs out of *Discharge Budget* but still have *Flexible Capacity*, *Ozone* allows the server to keep using green energy with reduced server frequency. If the stored energy drops below 60% of its installed capacity, *Ozone* will switch the load to the utility power side.

$P_{Load} > P_{Renewable}$				
1	R	DB > 0	upsC > 60%	Switch = 'N', cpuFreq = 'Highest'
2			upsC < 60%	Switch = 'Y', cpuFreq = 'TBD'
3		DB = 0	upsC > 60%	Switch = 'N', cpuFreq = 'Low'
4			upsC < 60%	Switch = 'Y', cpuFreq = 'TBD'
5	U	DB > 0	upsC > 80%	Switch = 'Y', cpuFreq = 'Highest'
6			upsC < 80%	Switch = 'N', cpuFreq = 'TBD'
7		DB = 0	upsC > 80%	Switch = 'Y', cpuFreq = 'Lowest'
8			upsC < 80%	Switch = 'N', cpuFreq = 'TBD'

$P_{Renewable} > P_{Load}$				
9	R	-	-	Switch = 'N', cpuFreq = 'Highest'
10	U	-	-	Switch = 'Y', cpuFreq = 'Highest'

***R** – Renewable power side; **U** – Utility power side; **DB** – Discharge budget; **upsC** – Estimated battery/ups capacity; **Y** – Switch the power supply; **N** – Don't switch; **TBD** – Server speed scales evenly depending on actual power budget* ;

Table 2: The power control decision tree of *Ozone*

5. EVALUATION METHODOLOGY

We develop an evaluation framework using our prototype system. The framework is configured into three layers: the *Power Budget Control Layer*, the *Oasis Operation Layer*, and the *Data Collection and Analytic Layer*.

In the *Power Budget Control Layer*, we feed the system with pre-defined power budget. We use the peak server power demand as default utility power budget of our system. To ensure fare comparison, we collect real-world solar power traces and use it as renewable power budget for all the experiments.

In the *Oasis Operation Layer*, we setup four 1U rack-mounted servers as computing load, as shown in Table 3. They are high-performance lower-power computing nodes that use Intel Core i7-2720QM 4-core CPU as processing engine. The measured idle power and peak power of each server are 21W and 55W, 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 live migrate the virtual machine by using command (*xm migrate DOMID IP -l*). Xen power management feature is also enabled to dynamically tune the vCPU frequency by using command (*xenpm set-scaling-speed cupid num and xenpm set-scaling-govornor cupid gov*). Our system kernel is configured with the on-demand frequency scaling governor. We set the minimum frequency as 0.8GHz and the normal frequency as 2.2 GHz.

We choose various datacenter workloads from Hibench [20] and CloudSuite [21]. 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. As shown in Table 4, we select ten workloads from five roughly classified categories. Within each experiment, a workload is executed iteratively.

In the *Data Collection and Analytic Layer*, we deploy front-end network server to communicate with the server cluster through a TP-Link 10/100M rack-mounted switch. The network server uses an AMD low power 8-core CPU with 16GB installed memory. We write system drivers using Linux socket to enable data communication between front-end server and the *Oasis node*. We store the collected battery charging/discharging statistics and the measured server power consumption data in a log file. We also use Watts UP Pro power meter [22] for our measurement in energy-related experiments. This power meter is able to display instantaneous power consumption with relatively high accuracy (1.5%). It also provides internal memory for storing up to 120K history power data.

We evaluated the battery lifetime based on the observed battery usage profile. We assume the battery has a cycle life of 5000 times and a maximal service life of ten years. Note that *Oasis* design is orthogonal to the actual battery specifications. It could optimize the service life of a variety of battery scenarios, thereby increases the overall cost-effectiveness of datacenter.

Computing Nodes	
CPU	Intel Core i7-2720QM, 4-core, 2.2G, TDP 45W
Memory	8 GB, registered
Disk	Seagate Barracuda 7200RPM, 500 GB
M/B	SuperMicro ITX Socket G2 Motherboard
Front-End Server	
CPU	AMD Opteron 4256 EE, 8-Core 2.5G, TDP 32W
Memory	16 GB, registered
Disk	Seagate Barracuda 7200RPM, 1000 GB
M/B	SuperMicro ATX Socket C32 Motherboard
Switch	TP-Link TL-SF1024 24-Port Unmanaged 10/100M

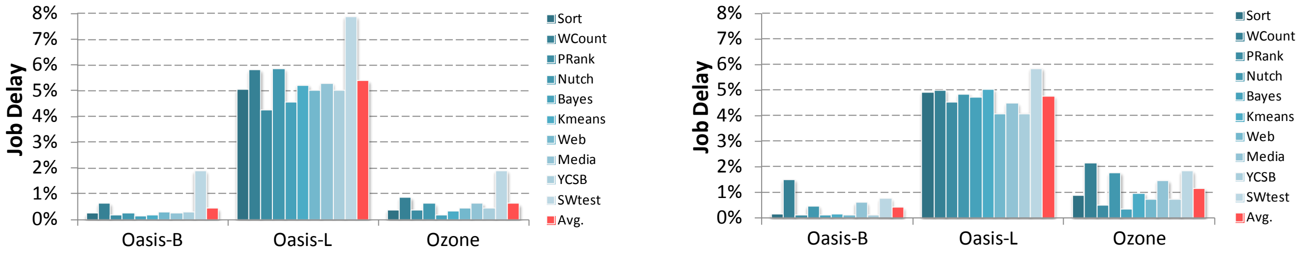
Table 3: Computing platform setup

Abbr.	**Workload**	**Category**
Sort	Sort program on Hadoop	Micro Benchmarks
WC	Word count program on Hadoop	Micro Benchmarks
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
Test	Software testing	Cloud Application

Table 4: Evaluated workloads [20, 21]

6. RESULTS

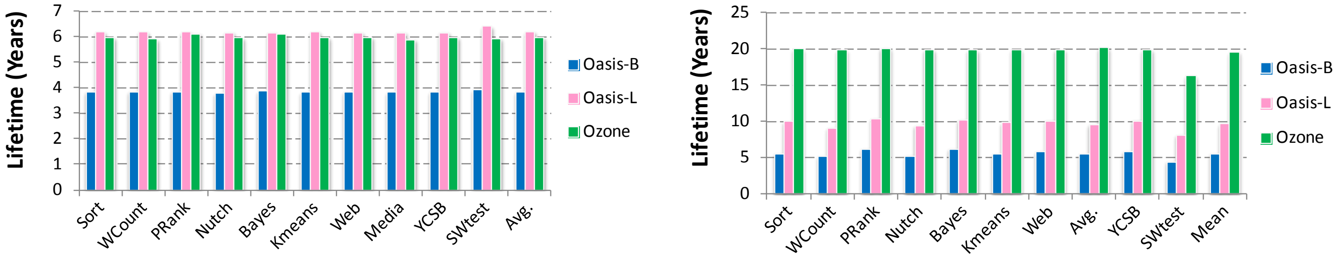
In this section we evaluate the impact of various power supply switching and server adaptation schemes. To be more specific, we evaluate three kinds of *Oasis* power management schemes: *Oasis-B, Oasis-L,* and *Ozone*. Their operation features are summarized in Table 5. In the following sub-sections, we first investigate the performance of various control schemes. We then discuss their impacts on battery lifetime and emergency handling capabilities. Afterwards, we evaluate the energy usage profile of different schemes. Finally, we discuss the cost issue of *Oasis*.



(a) High renewable power variability

(b) Low renewable power variability

Figure 9: The execution latency due to server performance scaling. *Ozone* could achieve very close performance compared to *Oasis-B*, which heavily uses battery to support high server speed.



(a) High renewable power variability

(b) Low renewable power variability

Figure 10: The estimated battery lifetime calculated from battery usage profile. Typically, *Ozone* shows much better battery lifetime compared to *Oasis-B*. In some scenarios when renewable power is constantly high, *Ozone* may result in under-utilization of battery system.

Schemes	Description
Oasis-B	Battery-oriented design. Focuses on using energy storage management. May result in low load performance if the stored energy is not enough.
Oasis-L	Load-oriented design. Focuses on the load power scaling capability of servers. It uses frequency scaling first, then leverage battery to compensate the power shortfall.
Ozone	Optimized control. Focuses on balanced usage of server load adaptation and the stored renewable energy.

Table 5: The evaluated *Oasis* power management schemes

6.1 Performance

Job turn-around time is a crucial metrics for emerging data-analytic workload in scale-out datacenters. Figure 9 shows the increase in job turn-around time under both high-variability and low-variability solar power generation scenarios. In Figure 9-(a), the mean turn-around time of *Oasis-B*, *Oasis-L* and *Ozone* is increased by 0.5%, 5.4%, 0.6%, respectively. In Figure 9-(b), the turn-around time of *Oasis-B*, *Oasis-L* and *Ozone* is increased by 0.4%, 4.7%, 1%, respectively.

As can be seen, *Oasis-B* shows the best performance. This is because *Oasis-B* trades off battery lifetime for server performance. In contrast, *Oasis-L* shows much higher performance degradation as it frequently lowers the processing frequency to match the inadequate renewable power budget. In Figures 9-(a) and 9-(b), the results of *Ozone* are very close to *Oasis-B*. On average, *Ozone* yields less than 1% performance degradation compared to *Oasis-B*, which heavily uses battery to provide power shortfall.

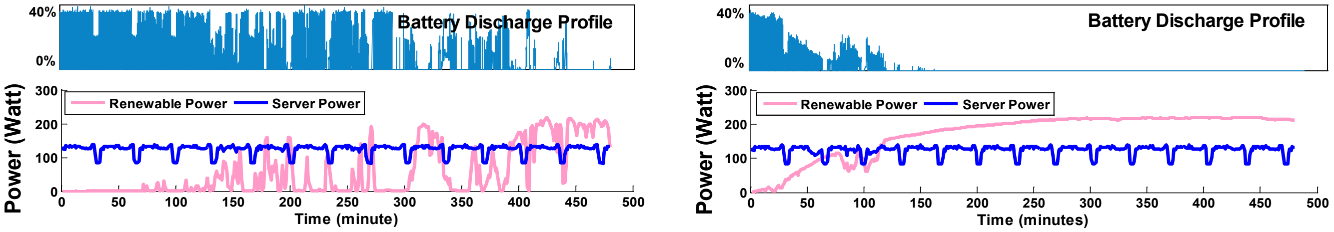
6.2 Battery Lifetime

In Figure 10 we estimate the battery service life based on detailed battery profiling information. Longer battery services life is favored as it lowers the total cost of ownership (TCO).

When renewable power supply varies significantly, the operation of server nodes typically requires substantial support from the energy storage. As a result, the predicted lifetime is much shorter than the designed service life (10 years), as shown in Figure 10-(a). On average, the lifetime of *Oasis-B, Oasis-L,* and *Ozone* is 3.9 years, 6.2 years, and 6.0 years, respectively. Due to the over-use of battery systems, the battery service life of *Oasis-B* is only 63% of *Oasis-L*. In contrast, the average battery lifetime of *Ozone* is 97% of *Oasis-L*. Compared to *Oasis-B*, *Ozone* could increase the lifetime by more than 50%.

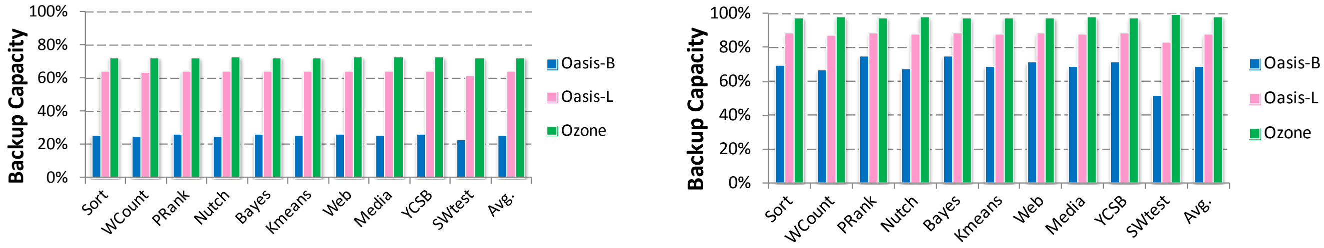
When renewable power supply becomes adequate, as shown in Figure 10-(b), the battery service of all the three power management schemes increases significantly. This happens because batteries are not discharged for most of time. However, commercial batteries cannot last for 20 years even if they are under-utilized. Many other issues such as aging and leakage will become the dominant factors for lifetime estimation if the batteries are used for an extended duration. In the real world, batteries are used much more frequently since renewable power system does not maintain peak output throughout its service life.

Figure 11 illustrates the problem of uneven battery usage. Note that when renewable output is constantly high, the distributed battery system is rarely used. In contrast, when renewable power fluctuates severely, batteries start to show heavy charging and discharging activities. Therefore, one can actually save plenty of *Discharging Budget* in scenarios like Figure 11-(b). This amount of *Discharging Budget* saving can be further used in scenarios like Figure 11-(a) to provide necessary load power support. In fact, the *Oasis* user can opportunistically leverage stored energy (even if current *Discharging Budget* is zero) to boost system performance without significantly affecting the battery service life. As long as the renewable power generation is likely to be adequate in the near future, the controller can advance server system certain amount of stored energy (if there is still *Flexible Capacity* in the battery).



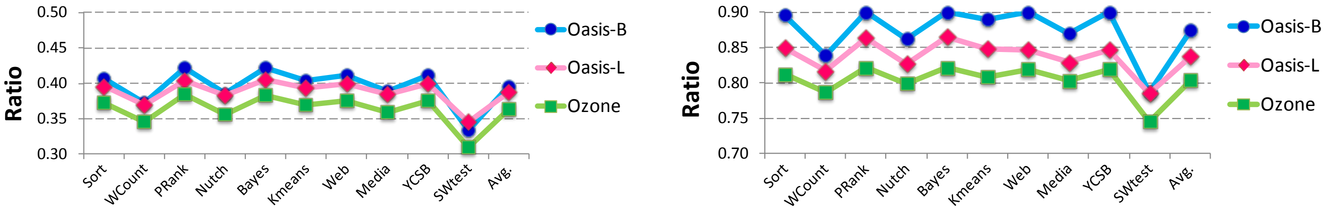
(a) High renewable power variability (b) Low renewable power variability

Figure 11: Battery discharging profile and the renewable power supply and server power demand traces (workload: Nutch page indexing).



(a) High renewable power variability (b) Low renewable power variability

Figure 12: The average battery backup capacity. *Ozone* maintains high backup time due to its better battery capacity management capability.



(a) High renewable power variability (b) Low renewable power variability

Figure 13: The ratio of green energy usage to overall power consumption. *Ozone* has relatively higher dependence on grid power.

6.3 Emergency Handling Capability

When datacenter servers scale out, their backup up power should be increased accordingly. While instantaneous workload performance boost is important for *Oasis* server, maintaining necessary backup capacity is even more important. A low battery backup capacity can pose significant risk since the backup generator may not be ready to pick up the load.

In Figure 12 we show the average UPS capacity during runtime for different power management mechanisms. *Ozone* could maintain around 73% backup capacity when renewable power variability is high and about 98% backup capacity when renewable power variability is low. *Oasis* suffers increased numbers of charging/discharging cycles in circumstances that renewable power varies significantly. Therefore, the backup capacity is low for all the three power management schemes in Figure 12-(a). Without setting a limit on the battery usage and the minimum battery stored capacity, the battery backup time can drops by 75% (i.e., *Oasis-B*).

6.4 Energy Usage Profile

Datacenters with heavy data-processing computing task often desire significant amount of runtime and consume large amount of energy. Leveraging green energy to provide additional power could save utility power bills considerably and lower the negative environmental impact of carbon-constrained datacenters. In Figure 13 we evaluate datacenter 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* yields 11% less renewable power rate when renewable power varies significantly and 8% less renewable power rate when renewable power generation is high. The reason *Oasis-B* shows high green energy usage is that it heavily uses battery to harvest renewable energy. In contrast, *Oasis-L* aggressively scales load power demand to match the renewable power generation.

7. COST ANALYSIS

Cost-effectiveness has become one of the top factors that drive server-class system design and optimization. In this section, we discuss the impact of *Oasis* on the cost of large datacenters. We estimate the system cost based on the collected operation log of *Oasis node*, trusted industry datasheets, manufacturer specifications, and government publications.

7.1 Cost Breakdown of Oasis Node

We first evaluate the cost of *Oasis* design (excluding the IT server cost and the cost of labor). Figure 14 presents two pie charts that show the cost breakdown. On the left we evaluate our *Oasis node* prototype. As can be seen, solar panel is the most expensive component (account for about 29% of the overall expenditure) in our renewable energy powered server system, followed by the PLC module (22%), the power inverter (16%), the battery (13%), and the HMI (9%).

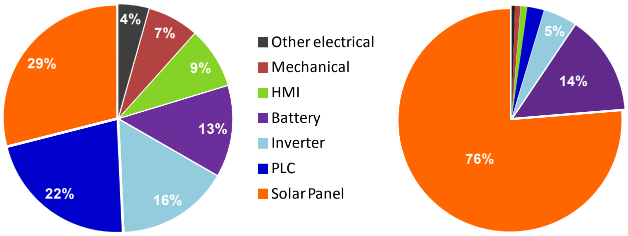


Figure 14: A cost breakdown of *Oasis*. Left: summary of our prototype. Right: estimation for a 5KW standard server rack. Some components such as HMI and PLC do not need to scale up when server system scales out. Solar panels, batteries, and inverters are dominant components.

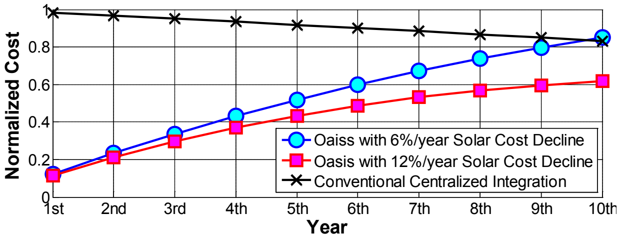


Figure 15: Cost trends of deploying *Oasis* for matching gradually rising computing demand in scale-out datacenters. In a typical 10-year datacenter life, gradually increase the capacity of installed renewable power system is more economical.

On the right chart we estimate the cost of *Oasis* design for a 40U standard server with moderate density (<10KW). We assume the rack is populated with 20 SuperMicro 1017R Xeon 1U servers. The cost of *Oasis* power control hub, including a 5KW inverter, is about $1.2K. It is approximately 4% of the total cost of the server equipment. While PLC and HMI are both major cost components in our prototype system, they actually account for only a very small portion of the total cost of ownership in large-scale datacenters. In addition, they do not need to scale up when more server systems are added. In contrast, power provisioning systems such as solar panels, batteries, and inverters need to increase their capacities to meet the growing load demand. Since *Oasis* is designed to take advantage of existing battery systems in a datacenter, solar panels and power inverters are often the major hardware additions.

7.2 Cost Projection for Large Deployment

Centralized renewable power integration has relatively low initial cost due to the scale effect. Recent report estimates that small-scale PV panel (around 5KW, 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 [23]. In addition, solar power inverter accounts for about 10%~20% of the initial system cost [24]. Central inverters in several MWs level are often cheaper compared to micro inverters (typically < 10KW) used in the *Oasis* controller. The former costs around $0.18/W, while the later costs around $0.5/W [25].

The main advantage of *Oasis* is that it allows users to gradually increase the installed renewable power capacity, thereby eliminating the inefficiency of over-provisioning. *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% per year, on average, from 1998~2011, and by 11%~14% from 2010~2011, depending on system size [25]. The cost of micro-inverter also decreases by 5% yearly [25].

Figure 15 illustrates how *Oasis* design helps to improve the overall cost-effectiveness of renewable energy powered scale-out datacenters. We assume that *Oasis* users evenly increase their deployment of *Oasis node* with a ten-year scale-out plan. (e.g., equip 10% 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% per year, and an optimistic decline rate of 12% per year. We calculate electricity cost savings of renewable energy powered systems based on real historical solar power traces. We use hourly solar irradiance measurement data (January 2003 ~ December 2012, 24 hours a day) provided by the NREL Solar Radiation Research Laboratory [26]. We assume the utility power is $0.1/kWh and datacenters can sell excess renewable power to the utility.

Figure 15 shows the cost overhead (total additional cost due to renewable power integration) for different design scenarios. The results are normalized to the one-time expenditure of datacenter-level renewable power integration. For conventional green datacenter that uses centralized renewable power integration, one can expect to get 17% investment return (due to the electricity cost savings) after ten years. However, this estimation is optimistic as the utility grid typically uses negotiated renewable power feed-in tariff that has a lower purchase price. In addition, for safety reasons, there is also a limit on the maximum amount of renewable power that can be synchronized with the utility power. The overall cost of *Oasis* is close to the conventional centralized design if solar power cost decreases by a conservative rate of 6% per year. If solar power cost declines faster, namely, 12% per year, *Oasis* could result in 25% less overhead cost, compared to a centralized design.

Note that there is no one-for-all universal design. Green datacenters with centralized solar power system may be a better choice if one has a firm goal of renewable power integration capacity and a confidence in future load power demand. However, *if one wants to enhance datacenter operational resiliency and seeks to gradually incorporate renewable energy to avoid over-committing capital, Oasis provides an attractive alternative.*

8. RELATED WORK

In this section we describe the state-of-the-art design of green computing systems that allow datacenters to leverage alternative power sources and reduce average server carbon footprint.

The scale-out model starts to draw great attention these days as emerging cloud application and data-processing workloads tend to scale well with large numbers of compute nodes. There has been several pioneering work that introduces processor and system level design methodologies for scale-out systems [27, 28]. At server cluster-level Kontorinis et.al [10] propose distributed UPS systems for more cost-effective power over-subscription. Recently, Wang et.al [29] investigate the power scarcity issue in datacenters and proposes power distribution virtualization techniques for managing power-constrained datacenters. Different from existing designs that emphasis improving server efficiency and density to free up datacenter power capacity, this work looks at approaches that provide additional power capacity incrementally to power-constrained servers to enable them to scale out.

We observe three interesting stages of development in the design and management of computing systems that takes advantage of green energy systems. At first, designers mainly focus on hardware and system control techniques, with an emphasis on adapting server power to the time-varying renewable power budget [30, 31]. Following that, the second stage features several more flexible solutions that leverage workload adaptation [32-35]. The main idea is to shift deferrable workloads to a time window in which renewable generation is sufficient (temporal adaptation), or to relocate workloads to a different system where power budget is abundant (spatial adaptation). In the third stage, the gap between power supply management and workload management starts to diminish. For example, recent designs have highlighted approaches that cooperatively tune both energy sources and workloads to achieve an optimal operating point [36, 37].

Existing work on carbon-aware systems can also be roughly classified into three categories, as discussed below.

Focusing on Supply-Load Matching

The dominate design pattern for managing mismatches between server power demand and power supply budget is to enable supply-following (a.k.a. supply-tracking) computing load to eliminate supply-demand mismatches. *SolarCore* [30] leverages per-core DVFS on multi-core systems to track the peak solar power budget while optimally assigns the power across different workload. *Blink* [31] leverages the on/off power cycles of server motherboard for tracking wind power supply. Their goal is to minimize the negative impact of temporary server shutdown on internet applications. Recently, *iSwitch* [38, 39] proposes handling renewable power budget through dynamic VM live migration between two clusters. It emphasizes different supply/load tuning policies for different renewable power scenario. In addition, *Chameleon* [40] proposes using online learning algorithm to dynamically select power supplies and power management policies. However, it mainly focuses on server level power control. Similar to *iSwitch*, [41] also divides datacenter clusters into a brown part (which uses utility power) and a green part (which uses renewable power). While this work proposes a power delivery architecture that is similar to that of *Oasis*, it assumes centralized batteries and integrates renewable power at cluster level.

In contrast to the supply-following based design, our recent work on load-following based design [37] takes advantage of the self-tuning capabilities of some renewable power supplies to match the changes in datacenter server load. In [37], we adjust datacenter power demand for improving load following efficiency.

There has been prior work exploring fine-grained renewable power integration in datacenters. For example, Deng et al. [42] investigate the use of grid-tie inverters for managing renewable power distribution. However, their work focuses on concentrating renewable power on green servers and does not consider the role of distributed batteries and modular renewable power supplies.

Several recent papers have discussed the role of battery in server clusters [43, 44]. These papers propose using energy storages devices to shave peak server power, manage demand-supply mismatch, and avoid unnecessary load migration.

Prior proposals typically assume that the interface between renewable power source and server system is ready. Although future smart grid is expected to feature smart communication gateway for providing connectivity and interactive control between onsite power generator and computing load, currently such interface is not widely adopted. In this study we build our supply-load cooperative control infrastructure and explore cross-layer power management schemes to achieve the best design tradeoff.

Focusing on Resource Planning

Many proposals focus on optimizing cost and energy utilization in green datacenters. For example, Liu et.al [35] model and evaluate dynamic load scheduling scheme in geographically distributed systems for reducing datacenter electricity prices. Zhang et.al [33] discuss cost-aware load balancing that maximizes renewable energy utilization. Deng et.al [34] explore algorisms for optimizing clean energy cost for green Web hosts. Recently, Ren et.al [45] demonstrate that intelligently leveraging renewable energy (self-generation or purchasing) can lower datacenter costs, not just cap carbon footprints. In this study we investigate power provisioning architecture and power management schemes to allow scale-out datacenters to maximize the benefit of green energy.

Investigating Field Deployment

Several studies have demonstrated the feasibility of renewable energy powered datacenters. These designs typically employ energy storage devices, grid-tie power controller, or a combination of both to manage renewable power. For example, HP Labs [46] tests a renewable energy powered micro-cluster called *Net-Zero* for minimizing the dependence on traditional utility grid. Their scheduling considers shedding non-critical workload to match the time-varying solar energy output.

The most similar prior work is *GreenSwitch*, a workload and energy source co-management scheme on a prototype green datacenter called *Parasol* [36]. In this work the authors highlights datacenter free cooling, low-power server nodes, renewable power prediction, and net-metering mechanism. Similar to [36], our work also faces the problem of solar power supply variability and datacenter power demand fluctuation. However, *Oasis* & *Ozone* differs from *Parasol* & *GreenSwitch* in five main aspects:

- First, our work focuses on incremental solar power integration at PDU level. While [36] targets a broad category of systems from warehouse-scale clusters to small server containers, its discussion mainly focuses on datacenter-level solar power integration and management.
- Second, *Oasis* does not synchronize solar power supply with the utility power and therefore is more reliant on energy storage devices. The advantage of such design is that it facilitates gradual renewable power expansion and allows heavily renewable energy penetration.
- Third, our system emphasizes the role of hardware power management. *Oasis* dynamically adjusts its load processing speed in response to the time-varying energy source conditions. In contrast, [36] emphasizes the role of model-based software power prediction. It uses workload characteristics to guide energy source switching.
- Forth, *Oasis* & *Ozone* pays more attention to battery capacity management and lifetime optimization, while [36] uses large batteries to lower electricity cost in the face of renewable power and utility grid net-metering.
- Fifth, our prototype system uses commodity Intel Core i7 series servers, while *Parasol* & *GreenSwitch* explores low-power Atom-based computing nodes.

9. CONCLUSIONS

In this paper we envision the long-term competitiveness of introducing modular green energy sources into datacenter capacity expansion plan. We present *Oasis*, a novel power provisioning architecture that enables power-/carbon- constrained datacenters to scale out. Our system leverages modular renewable power supplies and emerging distributed battery architecture to provide automated power provisioning and orchestration. *Oasis* naturally supports the power needs of scale-out systems and could enable datacenters to double its capacity with up to 25% less overhead cost.

We propose *Ozone*, a cross-layer power management scheme for *Optimized Oasis Operation (O_3)*. *Ozone* enables *Oasis* system to distill crucial runtime statistics of different power sources to avoid unbalanced usage of power source. Our results show that *Ozone* could help *Oasis* to reduce workload execution delay to 1%, extend battery lifetime by over 50%, and increase backup time by 1.9X, while still maintaining satisfactory green power usage rate.

Oasis helps to drive scalability, flexibility, sustainability and reliability in datacenters. Our work intends to show how IT design and operation can be carbon-conscious and social responsible. We expect that *Oasis* will encourage and facilitate innovative green computing research. This paper will also provide useful experiences and insights for both academics and practitioners who intend to build more efficient and cost-effective green computing systems.

ACKNOWLEDGEMENTS

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] C. Belady, Projecting Annual New Datacenter Construction Market Size, Global Foundation Services, Technical Report, 2011

[2] J. Koomey, Growth in Data Center Electricity Use 2005 to 2010, Analytics Press, 2011

[3] The 2012 Uptime Institute Data Center Industry Survey, The Uptime Institute, 2012

[4] D. Bouley, Estimating a Data Center’s Electrical Carbon Footprint, Schneider Electric White Paper Library, 2011

[5] http://www.microsoft.com/environment/

[6] http://www.ibm.com/ibm/environment/climate/

[7] http://www.google.com/green/energy/

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

[9] http://green.ebay.com/greenteam/ebay/

[10] 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, International Symposium on Computer Architecture (*ISCA*), 2012

[11] P. Sarti, Battery Cabinet Hardware v1.0, Open Compute Project, 2012

[12] H. Sher, and K. Addoweesh, Micro-inverters—Promising Solutions in Solar Photovoltaics, Energy for Sustainable Development, Elsevier, 2012

[13] B. Fortenbery, and W. Tschudi, DC Power for Improved Data Center Efficiency, Lawrence Berkeley National Laboratory, Technical Report, 2008

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

[15] Modbus Protocol, http://www.modbus.org/

[16] X. Fan, W. Weber, and L. Barroso, Power Provisioning for a Warehouse-Sized Computer, International Symposium on Computer Architecture (*ISCA*), 2007

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

[18] HOMER Energy Modeling Software, http://homerenergy.com/

[19] 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. Journal of Power Sources, 2006

[20] S. Huang, J. Huang, J. Dai, T. Xie, and B. Huang, The HiBench Benchmark Suite: Characterization of the MapReduce-Based Data Analysis. Data Engineering Workshops, IEEE International Conference on Data Engineering, 2010

[21] CloudSuite 2.0, http://parsa.epfl.ch/cloudsuite

[22] Watts Up? Meters, https://www.wattsupmeters.com

[23] D. Feldman, G. Barbose, R. Margolis, R. Wiser, N. Darghouth, and A Goodrich, Photovoltaic (PV) Pricing Trends: Historical, Recent, and Near-Term Projections, *Joint Technical Report*, National Renewable Energy Laboratory and Lawrence Berkeley National Laboratory, 2012

[24] A Review of PV Inverter Technology Cost and Performance Projections, Navigant Consulting Inc. and National Renewable Energy Lab, Technical Report, 2012

[25] R. Simpson, Levelized Cost of Energy from Residential to Large Scale PV, The Applied Power Electronics Conference and Exposition, 2012

[26] SRRL Baseline Measurement System, http://www.nrel.gov/midc/srrl_bms/

[27] P. Lotfi-Kamran, B. Grot, M. Ferdman, S. Volos, O. Kocberber, J. Picorel, A. Adileh, D. Jevdjic, S. Idgunji, E. Ozer, and B. Falsafi, Scale-Out Processors, International Symposium on Computer Architecture (*ISCA*), 2012

[28] S. Li, K. Lim, P. Faraboschi, J. Chang, P. Ranganathan, and N. Jouppi, System-Level Integrated Server Architectures for Scale-out Datacenters, International Symposium on Microarchitecture (*MICRO*), 2011.

[29] D. Wang, C. Ren, and A. Sivasubramaniam. Virtualizing Power Distribution in Datacenters, International Symposium on Computer Architecture (*ISCA*), 2013

[30] C. Li, W. Zhang, C. Cho, and T. Li, SolarCore: Solar Energy Driven Multi-core Architecture Power Management, International Symposium on High-Performance Computer Architecture (*HPCA*), 2011

[31] N. Sharma, S. Barker, D. Irwin, and P. Shenoy, Blink: Managing Server Clusters on Intermittent Power, International Conference on Architectural Support for Programming Languages and Operating Systems (*ASPLOS*), 2011

[32] I. Goiri, K. Le, T. Nguyen, J. Guitart, J. Torres, and R. Bianchini, GreenHadoop: Leveraging Green Energy in Data- Processing Frameworks, European Conference on Computer Systems (*EuroSys*), 2012

[33] Y. Zhang, Y. Wang, and X. Wang, GreenWare: Greening Cloud-Scale Data Centers to Maximize the Use of Renewable Energy, International Middleware Conference (*Middleware*), 2011

[34] N. Deng, C. Stewart, D. Gmach, M. Arlitt, and J. Kelley, Adaptive Green Hosting, International Conference on Autonomic Computing (*ICAC*), 2012

[35] Z. Liu, M. Lin, A. Wierman, S. Low, and L. Andrew, Greening Geographical Load Balancing, International Joint Conference on Modeling and Measurement of Computer Systems (*SIGMETRICS*), 2011

[36] I. Goiri, W. Katsak, K. Le, T. Nguyen, and R. Bianchini, Parasol and GreenSwitch: Managing Datacenters Powered by Renewable Energy, International Conference on Architectural Support for Programming Languages and Operating Systems (*ASPLOS*), 2013

[37] C. Li, R. Zhou, and T. Li, Enabling Distributed Generation Powered Sustainable High-Performance Data Center, International Symposium on High-Performance Computer Architecture (*HPCA*), 2013

[38] C. Li, A. Qouneh, and T. Li, iSwitch: Coordinating and Optimizing Renewabel Energy Powered Server Clusters, International Symposium on Computer Architecture (*ISCA*), 2012

[39] C. Li, A. Qouneh, and T. Li, Characterizing and Analyzing Renewable Energy Driven Data Centers, International Joint Conference on Modeling and Measurement of Computer Systems (*SIGMETRICS*), 2011

[40] 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, International Symposium on Low Power Electronics and Design (*ISLPED*), 2013

[41] M. Haque, K. Le, I. Goiri, R. Bianchini, and T. Nguyen, Providing Green SLAs in High Performance Computing Clouds. International Green Computing Conference (*IGCC*), 2013

[42] N. Deng, C. Stewart, and J. Li, Concentrating Renewable Energy in Grid-Tied Datacenters, International Symposium on Sustainable Systems and Technology (*ISSST*), 2011

[43] S. Govindan,D. Wang， A. Sivasubramaniam, and B. Urgaonkar, Leveraging Stored Energy for Handling Power Emergencies in Aggressively Provisioned Datacenters, Battery Emergency, International Conference on Architectural Support for Programming Languages and Operating Systems (*ASPLOS*), 2012

[44] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar, Benefits and Limitations of Tapping into Stored Energy for Datacenters, International Symposium on Computer Architecture (*ISCA*), 2011

[45] C. Ren, D. Wang, B. Urgaonkar, and A. Sivasubramaniam, Carbonaware Energy Capacity Planning for Datacenters, International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems (*MASCOTS*), 2012

[46] M. Arlitt et al., Towards the Design and Operation of Net-Zero Energy Data Centers, IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (*ITherm*), 2012