Towards Sustainable In-Situ Server Systems in the Big Data Era

*Chao Li[2,1], *Yang Hu[1], *Longjun Liu[3,1], Juncheng Gu[1], Mingcong Song[1], Xiaoyao Liang[2], Jingling Yuan[4], Tao Li[1]

[1]University of Florida, [2]Shanghai Jiao Tong University, [3]Xi'an Jiaotong University, [4]Wuhan University of Technology

Email: lichao@cs.sjtu.edu.cn, huyang.ece@ufl.edu, longjun.liu@stu.xjtu.edu.cn, taoli@ece.ufl.edu

Abstract

Recent years have seen an explosion of data volumes from a myriad of distributed sources such as ubiquitous cameras and various sensors. The challenges of analyzing these geographically dispersed datasets are increasing due to the significant data movement overhead, time-consuming data aggregation, and escalating energy needs. Rather than constantly move a tremendous amount of raw data to remote warehouse-scale computing systems for processing, it would be beneficial to leverage in-situ server systems (InS) to pre-process data, i.e., bringing computation to where the data is located.

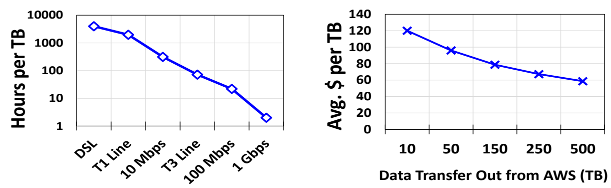
This paper takes the first step towards designing server clusters for data processing in the field. We investigate two representative in-situ workloads, which are normally generated from environmentally sensitive areas or remote places that lack established utility infrastructure. This very special operating environment of in-situ servers urges us to rethink conventional designs and explore standalone (i.e., off-grid) systems that offer the opportunity to benefit from local renewable energy sources. In this work we implement a heavily instrumented proof-of-concept prototype called InSURE: in-situ server systems using renewable energy. We develop a novel energy buffering mechanism and a unique joint spatio-temporal power management strategy to coordinate standalone power supplies and in-situ servers. We present detailed deployment experiences to quantify how our design fits with in-situ processing in the real world. Overall, InSURE yields 20%~60% improvements over a state-of-the-art baseline. It maintains impressive control effectiveness in under-provisioned environment and can economically scale along with the data processing needs. The proposed design is well complementary to today's grid-connected cloud data centers and provides competitive cost-effectiveness.

1. Introduction

Although many of the computing resources today are hosted in data centers, a tremendous amount of datasets are generated from distributed machines, monitors, meters, and various sensors. For example, there are approximately 30 million surveillance cameras deployed across the U.S., recording over 4 billion hours a week [1]. Even a single camera can create hundreds of gigabytes (GB) of data on a daily basis [2]. Similarly, smart sensors designed to monitor a wide area can easily generate several terabytes (TB) of data within a week [3]. Moreover,

ISCA'15, June 13-17, 2015, Portland, OR USA

* Authors with equal contribution



(a) Transfer time [6, 7] **(b) Transfer cost**

Figure 1: The overhead associated with bulk data movement

today's fast-growing scientific datasets (e.g., climate data and genome data) are typically distributed among many stations and research institutions around the world. Such wide-area collaboration on location-dependent data normally requires a routine data sharing of tens of petabytes (PB) every year [4]. According to a recent study by the Gartner Inc., transferring all these distributed datasets to a central location for processing will not be technically and economically viable in the big data era [5].

The enormous amount of datasets from distributed sources present significant challenges for data movement, especially when the volume and velocity of data are beyond the capability and capacity of today's commodity machines. Figure 1-(a) shows the data transfer time of 1 TB for typical network speed. Without high-throughput and scalable network, it could take days or weeks to move terabytes of data into the cloud [6, 7]. While the 10 Gigabit Ethernet-enabled equipment and emerging 40 Gigabit Ethernet are making their way into a data center's core backbone network, they are still not widely adopted at the network edge (i.e., near data source) due to high capital cost (CapEx) [8, 9]. As a result, Amazon Web Service (AWS) and Google Offline Disk Import now allow users to accelerate bulk data movement by shipping hard disks [10, 11]. Although some third-party solutions such as CERN's File Transfer Service [12] and LIGO's Data Replicator [13] could provide advanced data movement, they often require complex software and dedicated infrastructures, and therefore are only limited to very few scientific research communities [14].

In addition, the operating cost (OpEx) associated with data migration can quickly mount up. For example, Globus, a well-established bulk data sharing service provider, charges $1,950 per month for a 300 TB data transfer limit [15]. As of January 2014, Amazon charges over $60 for every 1 TB of data transferred out of its data centers, as shown in Figure 1-(b).

More importantly, for many data-driven projects that lack broadband access, the data movement issue becomes particularly acute. Some examples include oil/gas exploration [16], rural geographical surveying [17], astronomy observing in remote area [18], video surveillance for wildlife behavioral studies [19] and epidemic monitoring (e.g., Ebola) in Africa. While satellite/microwave based transmission has been used in some cases, it can cost over thousands of dollars per month with very limited network bandwidth [20].

Rather than constantly move a huge amount of data to a central data warehouse for processing, we instead explore a fundamentally different approach: tapping into *in-situ server systems (InS)*. The idea is to bring servers to where data is located to pre-process part, if not all, of the datasets. For instance, these servers can be used to eliminate duplicate copies, compress logs, or normalize data formats. Recently, a similar in-situ data processing scheme called *fog computing* has been proposed by Cisco to help prevent cloud systems from being overwhelmed [21]. However, it only uses Cisco's routers to process network traffic. The idea of in-situ computing has also been used in the HPC community to solve the I/O overhead for compute-intensive workloads [22, 23]. In this work we repurpose this concept to design server systems that can accelerate or facilitate the processing of distributed raw datasets

Our interest in in-situ server systems also arises out of the fact that modern data centers are heavily power-constrained, particularly when they employ power over-subscription to reduce cost. In the past five years, 70% companies have to build new data centers or significantly renovate existing facilities to handle the ever-growing traffic [24]. Meanwhile, recent studies are forced to aggressively discharge backup batteries to provision more servers under existing power budget [25-27]. As data continues to flood into data centers, it is not unusual that the utility power feeds are at full capacity and data centers do not have enough power to accommodate their growth [28, 29].

A significant challenge associated with in-situ processing is efficient power provisioning for servers running in the field. We find that ***a standalone wind/solar system with batteries as green energy buffer (e-Buffer) best suits the needs of in-situ servers and demands more attention*** for several reasons. First, conventional grid-tied designs may not be applicable since the construction and operation of transmission lines are often prohibitive in remote areas and hazardous locations. Even if the power line extension is technically feasible, grid-tied servers can violate environmental quality regulations in rural areas that are ecologically sensitive [30, 31]. In fact, to cap the significant IT carbon footprint, recent studies have already started to harness the power of green energy [32-39]. Further, in contrast to some other generators such as fuel cells and gas-turbines, wind and solar systems have many advantages such as absence of fuel delivery, easy maintenance, and less carbon emissions.

In this paper we present ***InSURE: in-situ server system using renewable energy***. As Figure 2 shows, we explore the opportunity to benefit from data pre-processing using a group of inexpensive, commodity servers that are placed near the data source. Specifically, we are primarily interested in in-situ datasets that need to be processed timely but do not have a very strong requirement for real-time processing. In fact, it has been shown that about 85% big data processing tasks can be deferred by a day [40]. Therefore, even if the renewable power output is intermittent and time-varying, we can still leverage it for processing many delay-tolerant data sets.

The main obstacle we face in developing InSURE is the lack of a cross-layer power management scheme that spans standalone power supplies and in-situ server systems. On the one hand, it is important to match the throughput of server clusters to the data processing demand. This allows InSURE users to timely process newly generated logs (so that geologists can use it to adjust their survey strategies) and to efficiently compress archival data (so that surveillance videos can be stored for extended periods). On the other hand, one must keep a watchful eye on the energy systems that directly support our servers. Without appropriate coordination, one may either lose the opportunity of harvesting enough renewable energy or incur unexpected power anomalies. Consequently, it can cause unnecessary data processing delay or even server shutdown.

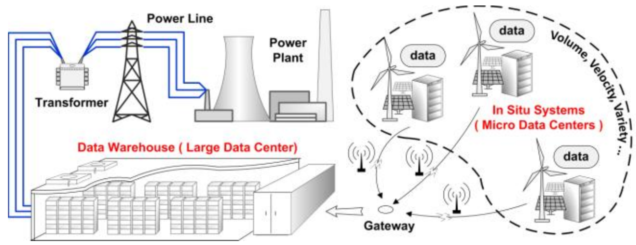


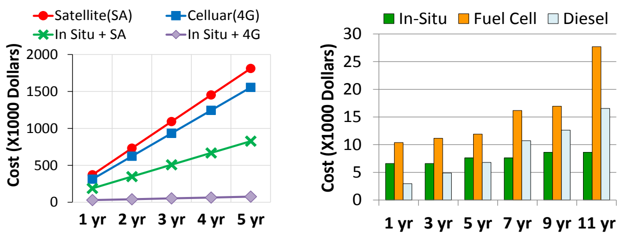
Figure 2: In-situ server system as an ancillary to future cloud

To overcome the above issue, we have developed a novel energy buffering mechanism and a unique joint spatio-temporal power management strategy that are tailored to the specific power behavior of standalone in-situ servers. They enable our system to intelligently reconfigure the size of energy buffer and accordingly adjust in-situ server loads during runtime. These two techniques provide several key benefits. First, they increase the overall efficiency of power delivery from standalone power supplies to in-situ server loads under varying renewable energy generation conditions. Second, they can greatly mitigate the frequency of server load shedding caused by various in-situ workload triggered demand-supply power mismatches. Third, they also balance the usage of different energy storage units and improve the longevity of our energy buffers.

We have implemented InSURE as a full-system prototype. It is a fusion of modular solar panels (1.6KW), a professionally assembled energy storage system, a Xeon-based micro server cluster, a software management platform built from scratch, and several other components such as internal communication infrastructure, power meters, and micro-controllers. Using our prototype and real in-situ workloads, we explore the technical and economic feasibility of in-situ data processing. We show that the proposed design is highly sustainable and is well complementary to cloud data center in the big data era.

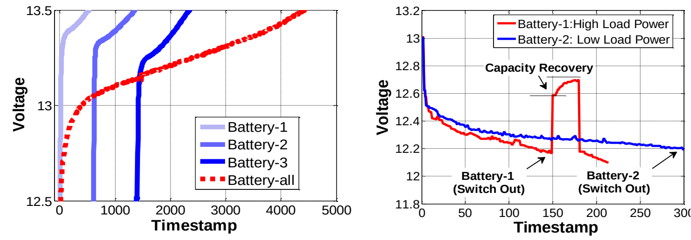
This paper makes the following key contributions:

- We explore *in-situ servers (InS)* for managing distributed big datasets today and tomorrow. We present *InSURE*, an *in-situ (standalone) server system using renewable energy*, and discuss its essential design considerations.
- We propose a novel energy buffering mechanism and a joint spatio-temporal power management scheme tailored to the behaviors of InSURE. It ensures highly efficient energy flow from the power supply to in-situ servers.
- We implement InSURE as a system prototype. We present detailed deployment experiences, demonstrate key design tradeoffs, and show that our optimizations can improve various measurement metrics by 20%~60%.
- We evaluate the cost benefits of InSURE. We show that it can economically scale along with different computing needs under various renewable energy availabilities.



(a) IT-related TCO (b) energy-related TCO

Figure 3: Cost benefits of deploying standalone InS



(a) individual vs. batch charging (b) high load vs. low load

Figure 4: Key properties of the energy buffer in standalone InS

The rest of this paper is organized as follows. Section 2 introduces InS. Section 3 proposes InSURE and its optimizations. Section 4 demonstrates our system prototype. Section 5 describes experimental methodology. Section 6 details deployment experiences and presents real-world case studies. Section 8 discusses related work and Section 9 concludes this paper.

2. In-Situ Standalone Systems: An Overview

The core idea of InS is to provide non-intrusive, eco-friendly data processing to minimize the overhead of bringing data to compute resource in the big data era. This section elaborates the concept of in-situ standalone systems and further motivates our design. We start by introducing typical in-situ applications and evaluating major cost issues. We then describe the properties of standalone energy systems with an emphasis on green energy buffers. Finally, we discuss the importance of smartly coordinating InS and standalone energy systems.

2.1 In-Situ Workloads and Cost Benefits

We investigate two representative types of application that may benefit from in-situ computing: *intermittent batch job* and *continuous data stream*. The former normally has large files that are generated periodically (often seen in engineering projects), while the latter faces constant influx of medium-sized data created by multiple machines (e.g., sensor data).

Oil Exploration (*intermittent batch job*): In oil and gas industry, massive volumes of micro-seismic data is collected and analyzed to guide the site selection and drilling [41]. An oil exploration project may involve tens of thousands of micro-seismic tests and each test can generate multiple terabytes of data [42]. Conventionally, these experiment data are processed at remote HPC cluster and usually rely on either expensive telecommunication transmission (e.g. via commercial satellite [20]) or time-consuming delivery via portable storage devices.

Video Surveillance (*continuous data stream*): Surveillance cameras are often deployed in hard-to-reach or hazardous areas to provide an understanding of wild life behaviors, volcano activities, and the source of local epidemics, etc. Many of these projects need a large volume of real-time and high-fidelity data which is far beyond the ability and capacity of conventional video monitoring systems. Conventional solution incurs huge human effort (e.g., manual data retrieval) and exposes researchers to hazard. It also incurs significant data storage overhead and time-consuming data aggregation.

We deploy 8 virtual machines (each VM has 4G memory and 2 virtual CPUs) on four HP ProLiant servers. We use open-source seismic data analysis software Madagascar [43] on 6 VM instances to conduct batch seismic data analysis. The in-situ workload is geographical surveying dataset for 225 square kilometers of real oil field [44]. We assume the seismic exploration happens twice a day and the data volume is 114GB per job. We also setup Hadoop based video analysis (pattern recognition) framework to process video stream data from 24 cameras (1280x720 resolution, 5fps). Details of our system prototype and configuration are discussed in Sections 4 and 5.

In both cases, processing data locally is much more cost-effective. Figure 3-(a) extrapolates the total computing cost (CapEx + OpEx) based on our real system prototype (detailed in Section 4). The satellite dish receiver costs about $11.5K and the service cost is $30K per month [45] or $0.14 per MB [20]. The hardware cost for cellular service is about $1K [46] and the service fee is $10 per GB [47]. The transmission cost can be several orders of magnitude larger than the cost of building our in-situ system prototype. In contrast to transferring all the data to remote data center via satellite, in-situ system can reduce over 55% operating cost if using satellite as backup communication method and 95% if using cellular service. It allows users to save over a million dollars in 5 years.

Onsite Generator	Energy-related CapEx	Energy-related OpEx
Diesel Generator	$370 per kW lifetime 5 yr	$0.4/kWh (diesel fuel price is $4/gallon)
Fuel Cells	$5/W, FC stack life 5yr, full system life 10yr	$0.16/kWh (natural gas is $14 per cubic ft.)
Solar + Battery	battery life 4 yr, 2$/Ah; solar panel 2$/W	N/A

Table 1: Parameters used in energy cost evaluation [48-53]

2.2 Standalone System and Energy Buffering

Whereas cloud data centers are grid-connected (mostly dual utility feeds), in-situ servers demand different power provisioning scheme. This is mainly because many data acquisition sites are temporary or difficult to reach - they lack established utility infrastructure. Thus, standalone power supplies such as solar/wind system (with commodity batteries as energy buffer) are often more suitable for data processing in field. They can provide eco-friendly energy without the fuel delivery needs like diesel generators or fuel cells do. Such green energy powered standalone systems are also economical. As shown in Figure 3-(b), fuel cell is still an expensive choice right now due to its relatively high initial CapEx. Although diesel generators have low CapEx and OpEx, they are not designed for supplying continuous power and often incur lifetime problems. The main OpEx of standalone solar system is the depreciation cost of energy storage (i.e., batteries). In this work our proposed power management scheme can actually extend their life.

Compute Capability	Avg. Pwr. (watts)	Availability	Throughput (GB/hour)
8VM (High)	1397	57%	**14.0**
4VM (Low)	696	100%	(Better) **16.5**

Table 2: Data throughput of seismic data analysis with the same energy budget (2kWh)

Compute Capability	Avg. Pwr. (watts)	Delay (minute)	Throughput (GB/hour)
8VM (High)	1411	0	(Better) **0.21**
6VM	1050	0.25	**0.17**
4VM	686	0.5	**0.10**
2VM (Low)	335	1.5	**0.07**

Table 3: Data throughput of Hadoop video analysis with the same energy budget (2kWh)

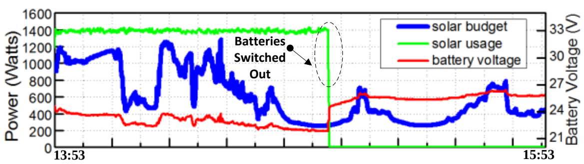


Figure 5: Snapshot of a 2-hour traces for seismic analysis

In off-grid situations, green energy buffers (e-Buffer) play a crucial role in maintaining high efficiency. During charging, for example, concentrating the limited green power budget on fewer batteries is often beneficial. This is because the charge acceptance rate of a near-empty battery is often much higher than a battery that is close to a full charge [54]. In Figure 4-(a), our real measurement shows that charging each battery unit one by one could reduce total charge time by nearly 50% compared to batch charging (i.e., charging all batteries simultaneously). In addition, during discharging, batteries incur super-fast capacity drop at high current. However, this temporary capacity loss can be recovered to a great extent during periods of very low power demand (known as "recovery effect" [55]), as shown in Figure 4-(b). Without careful management, the battery voltage drop can trigger emergency handling control and result in service disruption. Moreover, the aggregated electric charges (Ah) that flow through the e-Buffer is almost constant for a given battery unit before it wears out. This has been verified in extensive test on lead-acid batteries that undergo different charge/discharge regimes [56]. Therefore, one should also carefully balance the usage of the electric charge stored in every battery unit.

2.3 The Necessity for Cross-Layer Coordination

When in-situ workloads meet standalone power sources, it is the compute node and energy buffer that link them together. Therefore, it is crucial to judiciously manage both compute node and energy buffers. In fact, this can be very challenging.

First, the *intermittent batch job* and *continuous data stream* require different power management policies. Changing the number of VMs assigned to each job or adding other computing resources during job execution are difficult and in many cases impossible. In contrast, it is fairly easy to adjust the VM configuration during the period between two short time windows of the video streams. For some long-running batch jobs, increasing VM instances may not help improve productivity; on the contrary, our results show that it can degrade throughput by 15%, as shown in Table 2. The main reason is that the high

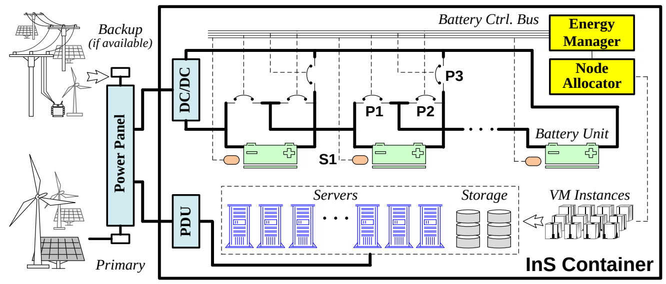


Figure 6: Schematic diagram of InS with adaptive energy buffer and smart node allocator. Although we focus on stand-alone system, it also supports a secondary power (if available).

server power demand can trigger increased number of checkpoints, causing undesirable service interruption (about 15 minutes for each server On/Off power cycle). In contrast, for video stream analysis workloads, a conservative system configuration (i.e., reduced VM instances) may not be wise. As shown in Table 3, reducing the number of active VM instances from 8 to 2 can reduce the data throughput by 66% and increase the service delay from zero to 1.5 minute per job.

In addition, conventional unified energy buffer lacks the ability to manage the energy flow from standalone systems to InS for two reasons. First, it has to be operated in either charging or discharging mode. The entire battery unit has to be disconnected from the load once its terminal voltage is below certain threshold for charging (or system protection reasons [55]). In this case InS has to be shut down and its solar energy utilization drops to zero. Figure 5 demonstrate this phenomenon on our prototype. Second, due to the very limited power budget in the in-situ environment, a unified energy buffer sometimes cannot receive the highest charging rate even if all the available solar power budget is used to charge the battery. Consequently, offline in-situ servers may incur extended waiting time.

3. Sustainable In-Situ Power Management

The unique operating environment of in-situ standalone servers requires a new, supply-load cooperative power management approach. In this work we propose InSURE, *in-situ server systems using renewable energy* (as the primary power source). The main goal of our design is to maintain highly productive data pre-processing and overcome the significant efficiency bottleneck caused by energy buffers. To achieve this, InSURE exploits two novel power management approaches:

1) **Reconfigurable distributed energy storage**
   We synergistically integrate a power switch network with distributed battery architecture. It allows the energy buffer to be operated in hybrid modes and adjust its size.
2) **Joint spatio-temporal power management**
   This technique jointly optimizes the efficiency of energy delivering 1) from standalone power supply to energy buffer and 2) from the energy buffer to compute servers.

3.1 System Overview

Figure 6 depicts the full system architecture of InSURE. A remarkable feature of InSURE is that it has built-in battery array that allows us to freely map a fraction of the stored green energy to servers. We leverage Facebook's external energy

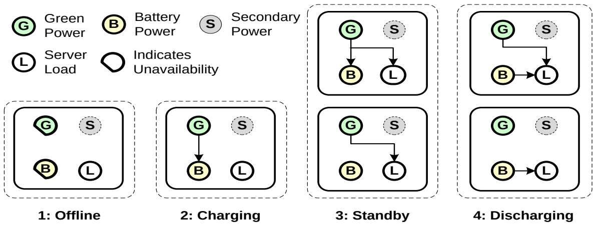


Figure 7: The various energy flow scenarios for InSURE.

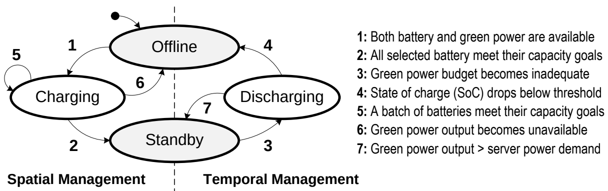


Figure 8: Operating mode transition of InSURE energy buffer.

storage architecture [57] and extend it by adding distributed power switches. This idea is partially enlightened by the reconfigurable battery design in the power system community [58]. As a result, InSURE can reconfigure the size of its energy buffer to optimize the energy flow from supply to load.

The major tuning knobs of our energy buffer are a set of switches managed by a PLC module. In Figure 6, three power switches (P1, P2, and P3) are used to manage the battery cabinets to provide different voltage outputs and ampere-hour ratings to servers. For example, if P1 and P3 are closed while P2 is open, the batteries are connected in parallel. If switches P1 and P3 are open while P2 is closed, the batteries are connected in serial. The PLC uses sensors (S1) to monitor the runtime status of each battery cabinet. It further communicates with the VM allocator to enable power-aware load matching.

In Figure 6, the in-situ servers and on-site power systems can actually be placed into a modular container. Some major components, such as the server rack, battery cabinet, and renewable power generator are all standardized and highly modular. As a result, the design complexity and maintenance cost is relatively low. If any of the above components requires replacement, the construction lead time is also very short.

3.2 Operating Mode

InSURE supports a variety of operating modes, as shown in Figure 7. Based on the state of the energy buffer, we categorize them into four types: *Offline*, *Charging*, *Standby*, and *Discharging*. In the *Offline* mode, batteries are disconnected from the server load for system protection purposes. In the *Charging* mode, onsite renewable power, if available, is used for charging batteries with the best achievable efficiency. Our design brings batteries online when they are charged to a pre-determined capacity (90%). In the *Standby* and *Discharging* modes, we use renewable energy (directly generated from onsite green generator or energy stored in the buffer) to power server clusters.

The transition between various operating modes is shown in Figure 8. Different battery units of InSURE's e-Buffer can be operated at different modes. They can adapt their operating modes to various scenarios based on the stored energy budget, server power demand, and battery health conditions.

Requires: The usage statistics *AhT[i]* of each battery unit *i*, designated total discharge D_L, desired lifetime T_L, and previously unused budget D_U

For each beginning of a coarse-grained control interval *T*
    Update the upper limit of battery discharge $\delta_D = D_U + D_L \cdot T/T_L$
For each battery cabinet *i* in the *offline group*
    Move *i* into *charging group* **If** $AhT[i] < \delta_D$

Figure 9: Spatial management scheme in the *Offline* mode. The goal is to avoid aggressive discharge and balance discharge.

Requires: The estimated green power budget P_G in the following battery charging period*;* the peak charging power P_{PC}

Calculate the optimal number for batch charging $N = P_G/P_{PC}$
While *charging group* has uncharged battery cabinets
    Select a batch of up to *N* cabinets C_N from *charging group*
    While current C_N capacity < 90% of its nominal capacity
        Continue charging
    Mark all batteries in B_N as charged
Move battery cabinets from *charging group* to *standby group*

Figure 10: Spatial management scheme in the *Charging* mode. The goal is to adapt the energy buffer size to renewable power budget to achieve fast-charging.

3.3 Spatial Management

The spatial power management scheme (SPM) fine-tunes the renewable energy harvesting (i.e., charging) process for in-situ server systems. It accelerates energy buffering and balances the usage of different battery cell to reduce wear-and-tear.

Other than treats all the battery units as a unified energy buffer, SPM focuses on selecting an optimal subset of the battery unit pool with a two-step control. First, the system selects battery units from the current energy storage pool based on the history usage record of each battery unit. Afterwards, it determines the optimal number of battery units for charging based on the available renewable power budget.

During runtime, InSURE maintains a battery discharge history table and monitors the state-of-charge (SoC) of each battery unit, as shown in Figure 9. At each time stamp *T*, the energy manger calculates a discharge threshold δ_D, which specifies the upper bound of the aggregated total discharge. Assuming that the lifetime discharge is D_L, the unused discharge budget in the last control period is D_U, and the desired battery lifetime is T_L, the discharge threshold is given by Eq-1.

$$\delta_d = D_U + {}^{T}\!/_{T_L} \cdot D_L \qquad \text{(Eq-1)}$$

We use the above threshold as the default criterion for determining whether a battery unit is over-used. Batteries are put into *offline group* if their aggregated discharge is greater than the threshold value. After the first screening, we obtain a group of battery that can be used in the incoming cycle.

In the second step, our system calculates the charging rate based on the available renewable power budget, as shown in Figure 10. If renewable power is inadequate, our system will reduce the number of battery units in the following round of batch charging. By concentrating the precious renewable power to fewer battery units, we could maintain a near-optimal charge rate in different renewable power generation levels. Once all the selected batteries are charged to a pre-determined level, they will be connected to the server cluster. In the *Standby* mode, batteries receive float charging.

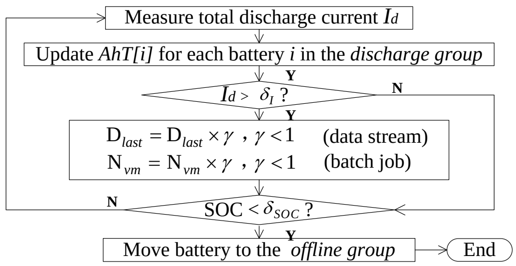


Figure 11: Flow chart of InSURE temporal power management

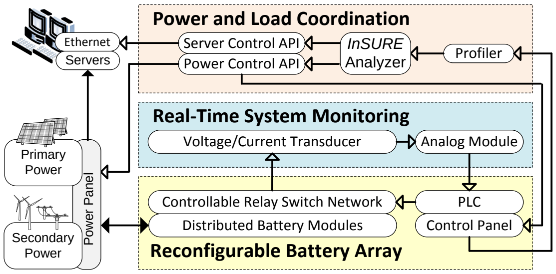


Figure 12: The structure of our verification platform

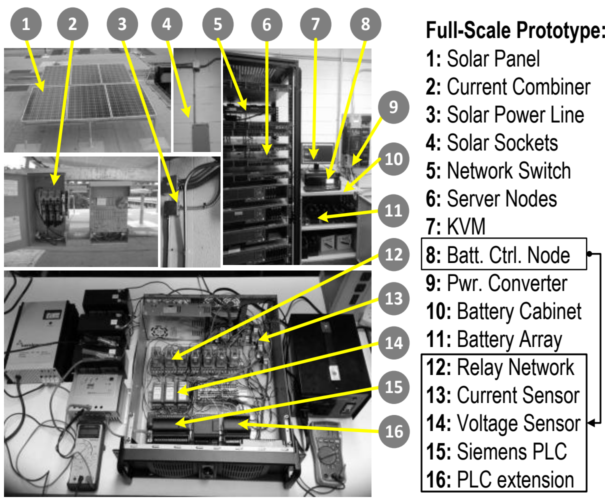


Figure 13: A full-system implementation of InSURE design

PROGRAMMABLE LOGIC CONTROLLER	Siemens S7-200 CPU224 PLC control module, 6ES7-214-1AD23-0XB0
PLC ANALOG INPUT MODULE	Siemens S7-200 6ES7-231-0HC22-0XA0, with 4 way analog signals input
NETWORK & COMMUNICATION INTERFACE	Cisco SRW2024 Gigabit Switch and Weintek MT8050i control panel
RECONFIGURABLE BATTERY ARRAY	Six UPG UB1280 12V 35AH batteries; Six IDEC RR2P 24VDC relays
BATTERY SENSOR	CR Magnetics CR5310 voltage transducer (In: 0-50V DC; Out: +/- 5V DC) HCS 20-10-AP-CL current transducer (In:+/-10A DC; Out: +/-4 V DC)
SOLAR POWER SYSTEM	Grape Solar PV panels. Total installed capacity 1.6KW

Table 4: Technical data of major hardware components

The SPM also offers the flexibility to temporarily trades off battery lifetime for better data processing throughput. It may gradually increase the battery discharge threshold instead of using a rigid value. If we use a fixed discharge threshold as calculated in Eq-1, it is possible that only a small number of batteries can be selected during a long period of high server power demand or low renewable power generation. By adding additional battery units to the already selected battery set, we can provide on-demand processing acceleration for a short period of time without significantly affecting battery life.

3.4 Temporal Management

To achieve continuous productivity, one needs to make the best use of the stored renewable energy. High-current discharging drains the battery quickly, but resulting in very limited energy delivery. Appropriate load power capping allows us to maintain a favorable amount of usable (online) battery units and avoids service disruption caused by data inrush. We use a temporal power management (TPM) scheme to improve the discharging effectiveness of batteries. The main idea is to allow the battery to partially recover its capacity during discharging period by reducing power demand. As shown in Figure 11, our system checks the server load level and the discharging current of online battery units at the beginning of each control period. If the discharge current is larger than a predefined threshold, our system will notify the server rack to cap power. For batch jobs, it will receive a duty cycle that specifies the percentage of time a server rack is allowed to run at full speed. Then the OS can use dynamic voltage and frequency scaling (DVFS) to adjust server speed based on the duty cycle. For data stream workloads that can be split into multiple small jobs, our system adjusts the number of VMs assigned to each job. In the meantime, InSURE also monitors the state-of-charge (SOC) of battery. When the battery units indicate low energy reserve, our system can temporarily shut down servers (VM states saved).

It is worth pointing out that the novelty of InSURE is not "distributed battery", but a new battery-aware energy flow management scheme for in-situ systems. Prior data center battery designs are only optimized under loose power budget constraints (i.e., with continuous utility power as backup) and does not consider the power variability in in-situ environment.

4. System Implementation

We have implemented InSURE as a three-tier hierarchical system, as shown in Figure 12. Its main functionalities are achieved through three modules built from scratch: (1) a reconfigurable battery array, (2) a real-time monitoring module, and (3) a supply-load power management node.

Figure 13 shows our full-system prototype of InSURE. We deploy four HP ProLiant rack-mounted servers (dual Xeon 3.2GHz processors with 16G RAM and 500G SAS HDD). The peak power demand of each server is around 450W and the idle power is about 280W. Table 4 summarizes the technical data of several major hardware components used in our system.

Reconfigurable Battery Array

Our customized battery system uses six 12V lead-acid batteries and a relay network. We use six 10A/24V DC relays as the power switches for reconfiguring the battery array. Each battery is managed independently using a pair of two relays (charging and discharging switch). These relays provide satisfactory mechanical life (10M cycles) and fast switching (25ms), and are ideal candidates for managing battery in our study.

We use a Siemens S7-200 CPU224 PLC module as the controller for our reconfigurable battery system. Its digital output is connected to the relay network and can energize or de-energize the coil of relay to perform battery switching.

Real-Time System Monitoring

The monitoring module detects the battery status and notifies the system whenever the battery configuration profile changes. It also collects key parameters such as charging and discharging current and battery terminal voltage.

To enable real-time monitoring, each battery in the system is equipped with a voltage transducer and a current transducer. Their outputs are further sampled by two analog input PLC extension modules that are coordinated by the central PLC. All the analog readings processed by the analog input module are stored in specific registers in the PLC.

We use an external control panel to link the battery system and the coordination node. The control panel communicates with the coordination server node via Ethernet using the Modbus TCP protocol [59], a widely used communication protocol for industrial electronic devices due to robustness and simplicity. We design corresponding encoding/decoding components to handle the data communication in our system.

Power and Load Coordination

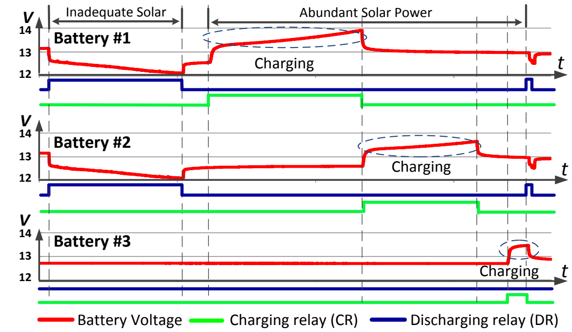
The top hierarchy of our design consists of a power and load coordination module. This module is implemented on a separate server node. It maintains runtime profiling data of the battery array and performs appropriate supply-load control.

We have designed a power supply switching API and a server control API. The former API provides necessary communication interface that allows the system to select its power source during runtime. The latter API is used to adjust server power demand through frequency scaling, server power state control, and virtual machine migration.

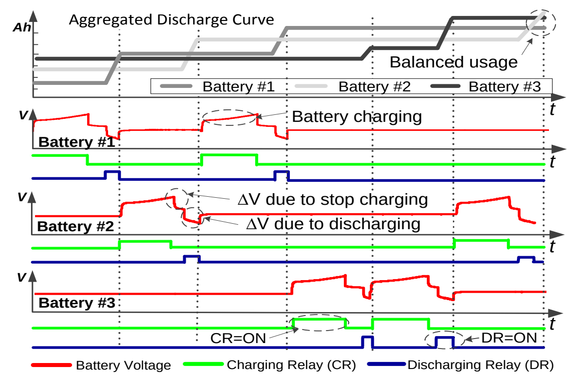
Power Behavior Demonstration

As mentioned earlier, the crux of our system is to maintain smooth and efficient energy delivery from standalone power system to energy buffers (e-Buffer) and finally to in-situ servers. Figures 14 (a) and (b) illustrate the system power behaviors using our recorded battery voltage traces and relay status logs.

Figure 14-(a) illustrates how our design performs timely solar energy harvesting. Initially, batteries #1 and #2 are both fully charged (*Standby* mode), whereas battery #3 is in low state of charge (SOC). As solar energy generation decreases, batteries #1 and #2 enter *Discharge* mode, and consequently their SOCs become lower than battery #3. At the time when InSURE receives adequate solar power, our controller starts to charge the battery array based on two principles: 1) give priority to low-SOC batteries if there are multiple batteries that have enough discharge budget, and 2) concentrate solar energy on fewer batteries for fast charging. In Figure 14-(b), our system selects batteries #1 and #2 since they have lower SOC (indicated by their terminal voltage). Our system starts to charge battery #3 after charging batteries #1 and #2 successively.



(a) from solar panel to e-Buffer (fast charging)



(b) from e-Buffer to InS (discharge balancing)

Figure 14: Demonstration of InSURE power behavior

Figure 14-(a) shows how our design manages the green energy flow from solar panel to e-Buffer with balanced battery usage. Initially, all three battery units are in *Offline* mode (discharged). When there is additional solar energy budget, our system selects batteries that have low aggregated total usage (Ah) for charging. Once reaching a pre-defined state of charge, the selected batteries will be put into *Standby* mode.

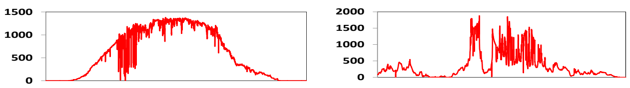
5. Experimental Methodology

We evaluate our design with both well-established micro benchmarks and real-world in-situ applications. Table 5 summarizes the workloads used in our experiments.

Micro Benchmarks: We use micro benchmarks to evaluate the power management effectiveness. We choose three benchmark programs from PARSEC [60], two from Hibench [61], and one from CloudSuite [62]. They cover a variety of in-situ data processing scenarios. For example, the *dedup* kernel represents data deduplication, which is the mainstream method to compress data; *vips* and *x264* are widely used image processing and video processing benchmarks; *wordcount* and *bayesian* are text-file processing programs (mimic the behaviors of log processing and analysis); the *graph* is a data mining application that uses a Twitter dataset with 11 million user data as input. Each workload is executed iteratively in our experiment.

	Workload	Input Size	Description
Micro benchmarks	dedup	672 M	Data deduplication
	graph	1.3 G	Graph analytics
	bayesian	2.4 G	Hadoop benchmark
	wordcount	1.0 G	Hadoop benchmark
	vips	2662x5500 pixels	Image processing
	x264	30fps/640x360 pixels	H.264 video encoding
In-Situ Apps.	seismic analysis	114G batch job; collected twice a day	Geo surveying data from 225 km^2 oil field
	video surveillance	5fps/1280x720 pixels 0.21G/min data rate	surveillance video generated from 24 cameras

Table 5: The evaluated benchmarks and in-situ workloads



(a) High solar generation (a) Low solar generation

Figure 15: Solar traces for evaluating micro benchmarks

In-Situ Applications: We also designed two case studies using representative in-situ applications on our prototype: 1) *seismic data analysis* (velocity analysis on 3D reflection seismic survey) widely used in the oil industry; and 2) *video surveillance analysis* (pattern recognition based video processing) for today's widespread surveillance cameras. The former workload is batch processing with a data rate of 114GB per job, two jobs a day. The latter is continuous data stream analysis based on videos generated from 24 cameras (0.21G/minute).

We host all workloads in virtual machines (VM) on Xen 4.1.2 hypervisor. Each physical machine (PM) hosts 2 VMs. Normally the first PM is turned on at 8:30AM, and the fourth PM is turned on 11:30AM. Starting from 4:00PM the first PM needs to be turned off and all PMs are shut down usually after 6:30PM. Our system automatically collects various log data and can initiate dynamic frequency scaling (DFS) on each PM. When solar power budget is inadequate, our system can further make checkpoint and all VM states are saved.

The time-varying nature of solar energy makes it difficult to compare different groups of experiments directly. Similar to [37], for micro benchmarks, we reproduce our experiments via collected real solar power traces and monitored workload runtime data. Note that we use this methodology only for comparing the optimization effectiveness of our spatio-temporal power management scheme with conventional designs.

As shown in Figure 15, we use two solar power traces that have different power variability patterns and average power generation levels. Our traces are collected from our roof-mounted solar panels that use a maximum power point tracking system to maximize its generation. We use daytime solar power traces collected from 7:00AM to 8:00 PM. The average power budget is 1114W for the high solar generation trace and 427W for the low solar generation trace. We use the dynamic solar power budget traces to precisely control our battery charger, so that the stored energy and the consumed green energy reflects the actual solar power supply across multiple experiments.

6. Results and Deployment Experiences

We start by analyzing the system behaviors of InSURE using real traces and system logs obtained from our prototype. We then evaluate our spatio-temporal power management scheme using micro-benchmarks. Finally we evaluate InSURE using real in-situ workloads and discuss its cost benefits.

6.1 System Trace Analysis

To understand the power behavior of inSURE. We investigate a typical system power trace collected from prototype, as shown in Figure 16-(a). We mark five typical regions (Regions A~D) that frequently appear in our everyday operation.

Our standalone in-situ system starts to charge a selected subset of batteries in the morning, as shown in Region-A. During this period, the battery voltage gradually increases until it reaches a preset value. Then the system enters standby mode.

Figure 16-(b) shows a zoomed view of a fraction of the system trace. As we can see, Region-B incurs a great deal of solar usage surges. This is because our system uses a *Perturb and Observe* (P&O) peak power tracking mechanism [63]. Our maximum power point tracker (MPPT) has built-in sensors that can identify if we have reached the optimal solar power output. To reach this point, the controller increases server load tentatively. This is reflected as many green peaks in Region-B.

Region-C shows the temporal control of our system. In this region, the load power demand is significantly larger than the maximum solar budget, and battery is in discharging mode. When our system realizes that the discharge current is unsafe, it triggers power capping (VM check-pointing and server shutdown in this case). As a result, the solar power demand of our in-situ system drops at the end of this region.

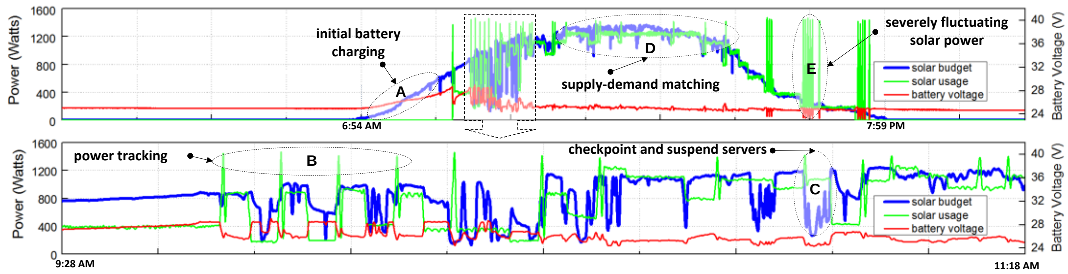
The Region-D is the most desirable region. Renewable power is adequate and system can harvest the maximum benefit from renewable energy powered data-preprocessing. In contrast, Region-E is an unfavorable region as severely fluctuating power budget can cause many supply-load power mismatches. Using peak power capping may solve this issue.

6.2 System Log Analysis

We further investigate InSURE by looking at the detailed system logs. In Table 6 we show a few key values that we extracted from three pairs of day-long operation logs (two sunny days, two cloudy days, and two rainy days). Each pair of traces has the same total solar energy budgets and very similar power variability patterns. We experiment with two power management schemes: 1) spatio-temporal optimization (Opt), and 2) aggressively using energy buffer (No-Opt).

Our results demonstrate a key trade-off in in-situ standalone server system design, i.e., the efficiency of the energy buffer can be improved at the expense of less renewable energy utilization. As Table 6 shows, the effective energy usage does not equal to the overall load energy consumption. This is because VM checkpointing operations and the on/off power cycles of servers consume large amount of energy but stall data processing progress. Since the optimized power management scheme (Opt) results in more VM operations and server on/off cycles, it yields lower effective energy usage, about 86% of a non-optimized scheme (No-Opt).

In fact, our spatio-temporal optimization trades off effective energy usage for good reasons. It improves the overall service life (the entire expected lifespan, typically 4 to 5 years for lead-acid batteries) of the e-Buffer and the average stored energy level by avoiding aggressive battery usage. As Table 6 shows, the standard deviation of battery terminal voltage of a non-optimized scheme is 12% higher than our design.



(a) Full-day operation demonstration of InSURE (Top) (b) Fluctuating power budget in a zoomed view (Bottom)

Figure 16: Solar power budget trace and battery state of charge (as indicated by voltage) collected using our monitors and sensors

		Load kWh Usage	Effective kWh Usage	Power Ctrl. Times	On/Off Cycles	VM Ctrl. Times	Minimum Battery Volt.	End of Day Voltage	Battery Volt. σ	Others in Common
Sunny (7.9 kWh)	Non-Opt.	6.7 kWh	6.4 kWh	12	8	8	23.6	25.2	1.05	Operating duration =11 hours; VM mgmt. overhead = 5 min; Battery initial voltage = 25.4V; Battery max voltage=28.8 V; Server pwr. consumption= 350 W; Max total server number = 4;
	Opt.	6.5 kWh	5.9 kWh	47	16	42	23.7	25.5	0.93	
Cloudy (5.9 kWh)	Non-Opt.	5.5 kWh	5.2 kWh	10	8	15	23.3	25.2	1.03	
	Opt.	5.0 kWh	4.2 kWh	51	20	51	23.2	25.3	0.92	
Rainy (3.0 kWh)	Non-Opt.	2.8 kWh	2.5 kWh	10	8	11	23.3	25.0	1.04	
	Opt.	2.6 kWh	2.1 kWh	33	15	38	23.3	25.4	0.93	

Table 6: Statistics of several key variables collected from the log. We show three typical solar power generation scenarios

Table 7 compares the impacts of integrating heterogeneous servers in our InSURE platform. We evaluate three types of workload: data deduplication, video processing, and data analysis application. An interesting observation is that, although emerging low-power servers (e.g. Intel Core i-7-2720 series) do not always guarantee the fastest immediate data processing speed, it contributes to significant energy efficiency improvement on our InSURE platform. They show better performance per watt and incur fewer On/Off power cycles (less overhead). Consequently, by using low-power servers, InSURE can improve data throughput by 5X~15X. In the future, we expect in-situ systems will benefit from more energy proportional designs being adopted by commodity server vendors.

	Data Size	Server Type	Exe. Time	Avg. Power	Data Processed per Unit of Energy per Node
dedup	2.6G	Xeon 3.2G	97s	360W	277G/kWh
		Core i-7	48s	46W	4.4T/kWh
x264	5.6M	Xeon 3.2G	4.6s	350W	12.4G/kWh
		Core i-7	4.7s	42W	101.3G/kWh
bayes	4.8G	Xeon 3.2G	439s	356W	111G/kWh
		Core i-7	662s	42W	621G/kWh

Table 7: Comparison of a legacy high performance server node to a state-of-the-art low-power server node.

6.3 Power Management Effectiveness

We first evaluate the power management effectiveness of InSURE using various micro benchmarks. Our baseline is a solar-powered in-situ system without the proposed spatio-temporal power management. Figures 17-19 show the results.

Our results demonstrate that InSURE can significantly improve the service availability of in-situ systems. Due to the optimized energy flow (from standalone systems to in-situ servers), InSURE shows 41% higher service availability under high solar generation. In Figure 17, when the solar energy generation is low, the improvement can reach 51%.

InSURE also saves the precious renewable energy stored in the e-Buffer throughout its operation. We refer to the average energy level of our e-Buffer as *energy availability*. In Figure 18, our system shows 41% more energy availability compared to our baseline. This can greatly improve the emergency handling capability of in-situ systems. The improvement is mainly a result from fast battery charging and smart load allocation that eliminates quick e-Buffer capacity drop.

We also expect a service life improvement of 21~24%, as shown in Figure 19. InSURE increases e-Buffer lifespan since it applies discharge capping and balancing across battery cabinets. Note that our optimization is conservative as we allow InSURE to occasionally use more stored energy than the predetermined threshold. By setting a more restrictive budget, one can further extend battery lifetime but may incur slight performance degradation. Exploring this tradeoff is our future work.

6.4 Full System Evaluation

We compare InSURE to a baseline in-situ design that adopts the power management approach of today's grid-connected green data centers [37, 38]. While our baseline system shaves peak power demand and tracks variable renewable energy, it can neither reconfigure its energy buffers nor adapt its nodes to the off-grid power supply. Figures 20 and 21 shows the results obtained from our prototype.

Overall, InSURE outperforms our baseline significantly (from 20% to over 60%) in terms of system uptime, data throughput, response time, energy availability, battery lifetime, and performance per Ah (defined as the total data processed per ampere-hour electricity flows through the e-Buffer).

A major observation is that the energy budget level could affect various optimization measurement metrics differently. In Figures 20 and 21, we broadly categorize these metrics into two types: *service-related metrics* and *system-related metrics*. The service-related metrics are more related with the user experiences and the system-related metrics mainly evaluates the energy efficiency and resiliency of in-situ systems.

The optimization effectiveness of InSURE on server-related metrics becomes greater when the solar energy is lower. In other words, the benefit of our joint spatio-temporal power management actually increases when the standalone in-situ system becomes heavily energy-constrained. The main reason behind this is that the charging process of our baseline system

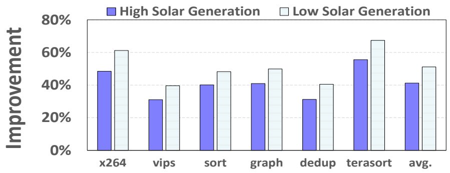


Figure 17: In-situ service availability

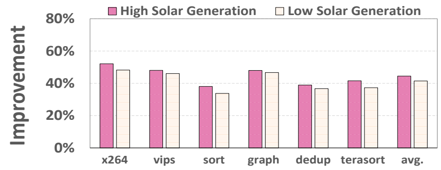


Figure 18: e-Buffer energy availability

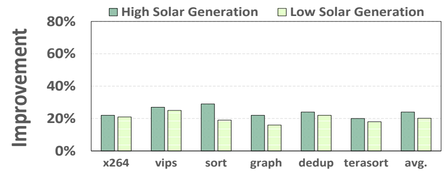


Figure 19: expected e-Buffer service life

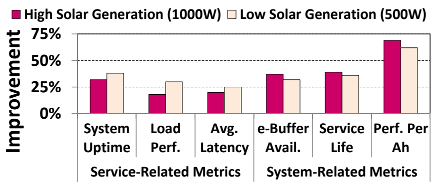


Figure 20: Results of in-situ batch job

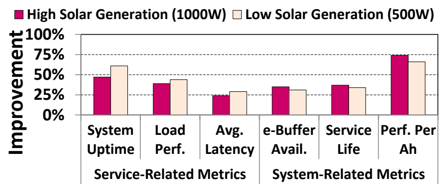


Figure 21: Results of in-situ data stream

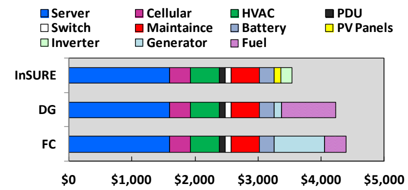


Figure 22: Annual depreciation cost

can be very lengthy when solar power budget is low. In addition, InSURE can reduce the significant energy overhead due to increased check points that are incurred by our baseline system.

When the solar energy is abundant, InSURE exhibits greater optimization results on system-related metrics. In fact, both InSURE and our baseline tend to deploy more compute instances and increase its charge and discharge frequency when the solar panel output is high. Due to a lack of intelligent discharge capping and balancing, the e-Buffer can be the major efficiency bottleneck of our baseline system.

The effectiveness of our temporal-spatial power management on system-related metrics is largely orthogonal to workload types. We observe that most of the performance statistics of *seismic data* are very close to *video surveillance*. In contrast, the service related metrics are normally workload dependent.

Our results also imply that InSURE could maintain its optimization effectiveness for in-situ systems that intend to oversubscribe their standalone power systems to reduce cost. In Figures 20 and 21, the height difference between the two bars (i.e., High Solar Generation and Low Solar Generation) is relatively small – less than 15% on average. Even if we cut the solar power budget in half in the Low Solar Generation scenario, it still shows very impressive overall improvement.

6.5 Cost Benefits

Although additional costs are added due to the inclusion of on-site renewable power supply, InSURE still provides competitive cost effectiveness. Take our full-system prototype (1.6kW) as an example, the solar array and inverter only account for 8% of the total annual depreciation cost, as shown in Figure 22. The cost of our e-Buffer (210Ah) is approximately 9% of InSURE. However, if powered by diesel generator, the same in-situ servers would require 20% additional equipment and fuel cost, not to mention that it increases the carbon footprint of in-situ servers. For fuel-cell based InS, it can generate carbon-neutral electricity with relatively cheaper fuels but the high capital cost of fuel cell stack has become the main design issue. Compared to InSURE, a fuel cell based InS can increase the cost by 24%. The maintenance cost is estimated to be 12% of InSURE. It is worth pointing out that the data transmission cost in rural areas can be several orders of magnitude larger than the cost of building InS. Adding hardware redundancy result in negligible cost increase but can further reduce maintenance cost. In addition, one can also leverage software fault-tolerant mechanisms to further reduce the maintenance cost.

Another advantage of InSURE is that it can economically scale along with different data processing needs, allowing for wide deployment. In places that have lower solar energy resources (e.g., indicated by sunshine fraction, the percentage of time when sunshine is recorded [64]), InSURE has decreased average throughput. In this case, one can scale out InSURE to meet the data processing demand. Although expending system capacity increases TCO, it is still much economical compared to sending unprocessed data to remote data centers. As shown in Figure 23, InSURE brings up to 60% cost savings.

The cost benefit of InSURE increases when local data generation rate increases. Figure 24 shows how the total cost of InSURE varies with data generation rate. There is a special point that the cost curve of InSURE interacts with cloud-based data processing. When the data generate rate is below this point (e.g., 0.9 GB/day for our prototype), our system exhibits higher operating cost compared to conventional cloud-based remote processing. If the data rate keeps increasing and researches 0.5 TB per day, our system could yield up to 96% cost reduction due to significantly reduced data transmission overhead.

We finally evaluate the cost savings of deploying InSURE in different in-situ big data scenarios [65-74]. We consider five application scenarios that have different data rates and deployment lengths, as shown in Figure 25. For some long-running data acquisition sites, we also consider the hardware replacement cost. Overall, InSURE provides an application-dependent cost saving rate ranging from 15% to 97%.

In-situ server clusters are complex systems and many other factors can affect their operating efficiency. For example, the intersection point in Figure 24 actually depends on the system capacity. Over-provisioning increases the TCO of InSURE and changes the position of the intersection point. Building efficient, cost-effective in-situ systems requires continued innovation in architecture and system design, which is our future work.

7. Related Work

Recent studies highlight various opportunities for building more powerful and efficient data centers, such as thermal control [75], deep sleep [76], peak power shaving [25], scale-out processors [77], market-driven system [78], hyperthread-aware power estimation [79], feedback control [80], and reconfigurable systems [81]. While all of the prior arts focus on pushing the efficiency limits of today's data center infrastructure, ***we instead explore the opportunity of offloading the ever-growing data processing burden to in-situ servers***. The importance of

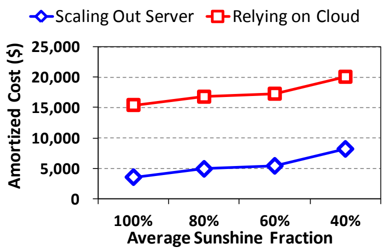


Figure 23: Total cost for scale-out based approach and cloud-based approach (i.e., sending data to cloud)

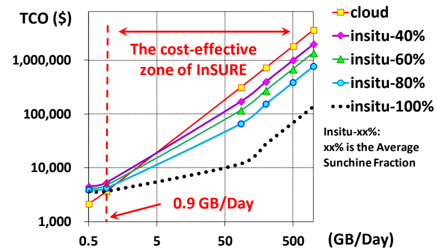


Figure 24: Cost comparison of cloud-based and in-situ computing under different data rate and power budget

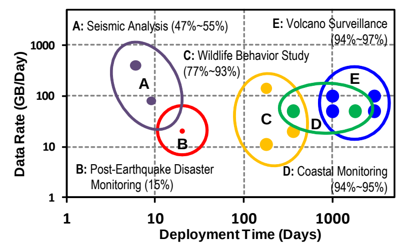


Figure 25: Application-specific cost analysis. The cost saving is represented by the size of the bubbles

developing such a new data-centric system that bridges the gap between computing capability and data growth has been recognized in different ways [82-84].

In-Situ Computing: The concept of in-situ computing has been proposed in the HPC and system research community to solve the I/O overhead issue. Several studies propose performing data analysis while scientific applications producing data [22, 23, 85] or moving computation from compute node to storage servers [86]. However, they only look at in-situ computing within the data center. They can neither solve the bulk data movement issue, nor address the grand power budget challenge faced by today's data centers. Several networking research [41, 87] also try to bring computation to the network edge near data sources. However, these proposals are limited to sensor nodes and network routers, lacking the necessary storage capacity and computing capability for handling massive amount of data.

Green Data Centers: Many recent work has explored data center powered by renewable energy [32-39, 88-91]. The most representative works are Parasol [37], Oasis [38], Blink [88], and Net-Zero [89]. However, our design differs from prior studies in both architecture and power management strategies.

- *Parasol*. The Parasol project includes a group of system design [33, 37, 90]. Its prototype is a solar-powered micro-data center backed by grid-tie and batteries. Its main feature is to smartly schedule deferrable jobs and select the source of energy to use. Our work distinguishes itself from Parasol in three aspects: (1) Parasol mainly focuses on data center level design, whereas InSURE looks at small-scale clusters deployed near the data. (2) The energy source selection strategy of Parasol is not applicable for standalone in-situ servers that have no access to utility grid. (3) Parasol is mainly concerned with renewable power variability, while InSURE mainly focus on the efficiency of energy delivery from standalone systems to in-situ nodes.
- *Oasis*. The highlight of Oasis is that it exploits incremental green energy integration at the PDU level for scaling out server clusters. It focuses on adding server racks to existing data centers and therefore is expected to incur the same data movement problem as cloud data centers. In addition, similar to Parasol, Oasis is a grid-connected system that relies on a controller to change power supplies.
- *Blink*. Blink leverages fast power state switching to match server power demand to intermittent power budget. The proposed design mainly focuses on internet workloads and lacks the ability to optimize energy flow efficiency.
- *Net-Zero*. Net-Zero is a solar energy powered server rack that matches load energy consumption to renewable energy generation to achieve carbon-neutral computing. It also relies on net-metering (a grid-dependent power synchronization mechanism) and cannot be used on in-situ systems.

Energy Storage Management: Batteries have attracted considerable attentions recently due to their importance in both small mobile systems [92] and large data centers [36, 25, 26, 93-95]. In contrast to prior energy storage systems designed for emergency handling purpose (rarely used) and peak shaving purpose (occasionally used), batteries used for standalone InS often incur *cyclic usage*, i.e., they are discharged in a much more frequent and irregular manner. In addition, prior studies overlooked several critical battery properties, resulting in sub-optimal tradeoffs for in-situ systems. In [96], the authors investigate a dynamic control scheme for distributed batteries, but it does not consider renewable energy and in-situ environment.

8. Conclusions

In this study we explore pre-processing of data generated in the field. Specifically, we find that in-situ standalone server system that is powered by renewable energy and backed by green energy buffers is especially promising. We show that efficient energy flow from standalone power supplies to energy buffer and finally to compute nodes is the crux of designing such an in-situ computing facility. This paper, for the first time, demonstrates the full-system implementation and a novel power management scheme for in-situ standalone server systems. Our system can bring 20~60% performance improvements in terms of system uptime, data throughput, energy availability, and battery lifetime. We believe in-situ systems provide a technically and economically viable way of tackling the incoming data explosion challenge. It will essentially open a door for a new class of sustainable computing in a world of ubiquitous data.

ACKNOWLEDGMENT

This work is supported in part by NSF grants 1423090, 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, and by three IBM Faculty Awards. Jingling Yuan is supported by NSFC grant 61303029. Chao Li is also supported in part by Yahoo! Key Scientific Challenges Program Award, Facebook Fellowship, and SJTU-MSRA Faculty Award.

References

[1] J. Vlahos. Surveillance Society: New High-Tech Cameras Are Watching You. *Popular Mechanics*, 2009

[2] Video Surveillance Storage: How Much Is Enough. White Paper, Seagate Technology, 2012

[3] Information Optimization: Harness the Power of Big Data. White Paper, Hewlett-Packard Company, 2012

[4] W. Johnston, E. Dart, M. Ernst, and B. Tierney. Enabling High Throughput in Widely Distributed Data Management and Analysis Systems: Lessons from the LHC. Technical Report, Lawrence Berkeley National Laboratory, 2013

[5] Gartner Says the Internet of Things Will Transform the Data Center. http://www.gartner.com/newsroom/id/2684616

[6] J. Gray, W. Chong, T. Barclay, A. Szalay, and J. Vandenberg. TeraScale SneakerNet: Using Inexpensive Disks for Backup, Archiving, and Data Exchange. Technical Report, Microsoft Research, 2002

[7] W. Vogels. Expanding the Cloud: Moving Large Data Sets into Amazon S3 with AWS Import/Export. Technical Blog, 2009, http://www.allthingsdistributed.com/

[8] Driving 10 Gigabit Ethernet Adoption in the Data Center. White Paper, Intel, 2009

[9] G. Chanda. The Market Need for 40 Gigabit Ethernet. White Paper, Cisco, 2012

[10] Amazon Import/Export.
http://aws.amazon.com/importexport/

[11] Offline Disk Import.
https://developers.google.com/storage/docs/early-access

[12] CERN's File Transfer Service.
http://information-technology.web.cern.ch/services/file-transfer

[13] LIGO Data Replicator
http://www.lsc-group.phys.uwm.edu/LDR/overview.html

[14] B. Allen, J. Bresnahan, L. Childers, I. Foster, G. Kandaswamy, R. Kettimuthu, J. Kordas, M. Link, S. Martin, K. Pickett, and S. Tuecke. Globus Online: Radical Simplification of Data Movement via SaaS. Technical Report, Argonne National Laboratory, 2012

[15] Globus Provider Plans
https://www.globus.org/providers/provider-plans

[16] Offshore Oil and Gas Supply. Working Document of the National Petroleum Council, 2011

[17] The Changing Geospatial Landscape. A Report of the National Geospatial Advisory Committee, 2009

[18] How Big Data Is Changing Astronomy (Again). *The Atlantic*, 2012

[19] W. Cox, M. Pruett, T. Benson, S. Chiavacci, and F. Thompson III. Development of Camera Technology for Monitoring Nests. USGS *Northern Prairie Wildlife Research Center*, 2012

[20] http://www.groundcontrol.com/Oil-And-Gas_Satellite.htm

[21] Cisco Technology Radar Trends. White Paper, Cisco, 2014

[22] F. Zhang, S. Lasluisa, T. Jin, I. Rodero, H. Bui, and M. Parashar. In-situ Feature-based Objects Tracking for Large-Scale Scientific Simulations. 2012 *SC Companion: High Performance Computing, Networking Storage and Analysis* (SCC), 2012

[23] S. Lakshminarasimhan, D. Boyuka, S. Pendse, X. Zou, J. Jenkins, V. Vishwanath, M. Papka, and N. Samatova. Scalable In-situ Scientific Data Encoding for Analytical Query Processing. *ACM International Symposium on High-Performance Parallel and Distributed Computing* (HPDC), 2013

[24] 2013 Data Center Industry Survey. *the Uptime Institute*, 2013
http://uptimeinstitute.com/2013-survey-results

[25] 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

[26] B. Aksanli, P. Eddie, and R. Tajana. Architecting Efficient Peak Power Shaving Using Batteries in Data Centers. *IEEE International Symposium on Modelling, Analysis & Simulation of Computer and Telecommunication Systems* (MASCOTS), 2013

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

[28] M. Ebbers, A. Galea, M. Schaefer, M. Khiem. The Green Data Center: Steps for the Journey. White Paper, IBM, 2008

[29] Z. Kishimoto, IT Controls Power at Data Centers, Technical Blog, AltaTerra Research Network
http://www.altaterra.net/blogpost/587092/168544/IT-Controls-Power-at-Data-Centers--part-1

[30] Environmental Impacts of Transmission Lines.
Public Service Commission of Wisconsin
http://psc.wi.gov/thelibrary/publications/electric/electric10.pdf

[31] Power, Pollution and the Internet.
http://www.nytimes.com/2012/09/23/technology/data-centers-waste-vast-amounts-of-energy-belying-industry-image.html?pagewanted=all

[32] K. Le, O. Bilgir, R. Bianchini, M. Martonosi, and T. Nguyen. Capping the Brown Energy Consumption of Internet Services at Low Cost. *International Green Computing Conference* (IGCC), 2010.

[33] Í Goiri, R. Beauchea, K. Le, T. Nguyen, M. Haque, J. Guitart, J. Torres, and R. Bianchini. GreenSlot: Scheduling Energy Consumption in Green Datacenters. *Supercomputing* (SC), 2011

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

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

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

[37] 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

[38] C. Li, Y. Hu, R. Zhou, M. Liu, L. Liu, J. Yuan, and T. Li. Enabling Datacenter Servers to Scale Out Economically and Sustainably. *International Symposium on Microarchitecture* (MICRO), 2013

[39] J. Berral, I. Goiri, T. Nguyen, R. Gavalda, J. Torres, and R. Bianchini. Building Green Cloud Services at Low Cost. *IEEE International Conference on Distributed Computing Systems* (ICDCS), 2014

[40] C. Bange, T. Grosser, N. Janoschek. Big Data Survey Europe: Usage, Technology and Budgets in European Best-Practice Companies. White Paper, BARC, 2013

[41] L. Shi, W. Song, Y. Xie, Z. Peng, and J. Lees. InsightTomo: In-Situ Seismic Tomographic Imaging System. *The ACM Conference on Embedded Networked Sensor Systems* (SenSys), 2014

[42] Drilling for Fuel. Scientific Computing World. 2009

[43] S. Fomel, P. Sava, I. Vlad, Y. Liu, and V. Bashkardin. Madagascar: Open-source Software Project for Multidimensional Data Analysis and Reproducible Computational Experiments. *Journal of Open Research Software* , 2013

[44] S. Fomel and B. William. Azimuthally anisotropic 3D velocity continuation. *International Journal of Geophysics*, 2011

[45] M. Ryan. The Need for Offshore Speed. Houston Business Journal, 2012
http://www.harriscaprock.com/downloads/HouBisJournal_12-12.pdf

[46] http://avalanwireless.com/shop/aw-netdrop-4g-cellular-ethernet-gateway/

[47] http://www.att.com/shop/wireless/plans/dataplans.html

[48] Costs of Utility Distributed Generators, 1-10MW.
http://www.publicpower.org/files/deed/finalreportcostsofutilitydistributedgenerators.pdf

[49] http://www.allworlddieselgen.com/faq.htm

[50] A. Riekstin, S. James, A. Kansal, J Liu, and E. Peterson. No More Electrical Infrastructure: Towards Fuel Cell Powered Data Centers. *Workshop on Power-Aware Computing and Systems* (HotPower), 2013

[51] Fuel Cell Technology: a Clean, Reliable Source of Stationary Power.
http://www.cesa.org/assets/Uploads/Resources-pre-8-16/CESA-fuelcelltechnology-may2010.pdf

[52] Short-Term Energy Outlook, Natural Gas.
http://www.eia.gov/forecasts/steo/report/natgas.cfm

[53] S. McCluer. Battery Technology for Data Centers and Network Rooms: Lead-Acid Battery Options. APC White Paper #30, 2011

[54] D. Linder. Handbook of Batteries. Third Edition, McGraw-Hill, 2002

[55] H. Kim, and K. Shin. Scheduling of Battery Charge, Discharge, and Rest. *IEEE Real-Time Systems Symposium* (RTSS), 2009

[56] K. Yabuta, T. Matsushita, and T. Tsujikawa. Examination of the Cycle Life of Valve Regulated Lead Acid Batteries. *International Telecommunications Energy Conference* (INTELEC) 2007

[57] P. Sarti. Battery Cabinet Hardware v1.0, Open Compute Project. 2012 http://www.opencompute.org/

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

[59] Modbus Communication Protocol Suite. http://www.modbus.org/

[60] http://parsec.cs.princeton.edu/

[61] S. Huang, J. Huang, Y. Liu, L. Li, and J. Dai, HiBench: A Representative and Comprehensive Hadoop Benchmark Suite, *IEEE 26th International Conference on Data Engineering Workshops* (ICDEW), 2010

[62] CloudSuite Benchmark, http://parsa.epfl.ch/cloudsuite

[63] N. Femia, G. Petrone, G. Spagnuolo, and M. Vitelli. Optimization of Perturb and Observe Maximum Power Point Tracking Method. *IEEE Transactions on Power Electronics*, Vol. 20, No. 4, July 2005.

[64] L. Wang and W. Chow. Solar Radiation Model. *Applied Energy*, Vol 69, 2011, pp 191-224

[65] http://www.oilandgaslawyerblog.com/2009/04/how-do-seismic-surveys-work.html

[66] http://switchboard.nrdc.org/blogs/tkiekow/the_true_cost_of_seismic_surve.html

[67] Land Seismic Surveys Will Never Be The Same, http://www.geospace.com/wp-content/uploads/2014/07/OYOatWorkNov08.pdf

[68] D. Larios, C. Rodriguez, J. Barbancho, M. Baena, F. Simon, J. Marin, C. Leon, and J. Bustamante. Intelligent Monitoring System for Bird Behavior Study. *E-Business and Telecommunications*, Vol. 455, 2014

[69] AJ Garcia-Sanchez, F. Garcia-Sanchez, F. Losilla, P. Kulakowski, J. Garcia-Haro, A. Rodriguez, JV Lopez-Bao, and F. Palomares. Wireless Sensor Network Deployment for Monitoring Wildlife Passages. *Sensors*, 10(8), 2010

[70] Using Sound to Monitor Rare and Elusive Species, Report, University of Alberta Biodiversity Monitoring Institute, 2014

[71] M. Nieto, B. Garau, S. Balle, G. Simarro, G. Zarruk, A. Ortiz, J. Tintore, A. Alvarez-Ellacuria, L. Gomez-Pujol, and A. Orfila. An Open Source, Low Cost Video-Based Coastal Monitoring System. *Earth Surface Processes and Landforms*, Vol 35, Issue 14, 2010

[72] B. Ando and E. Pecora. An Advanced Video-Based System for Monitoring Active Volcanoes. *Computers & Geosciences*, 32(1), 2006

[73] Broadband wireless communication system for emergency. http://english.sim.cas.cn/rh/as/200906/t20090629_9414.html

[74] Oil exploration. http://www.nscc-tj.gov.cn/en/example/example_1.asp

[75] W. Huang, M. Allen-Ware, J. Carter, E. Elnozahy, H. Hamann, T. Keller, C. Lefurgy, J. Li, K. Rajamani, and J. Rubio. TAPO: Thermal-Aware Power Optimization Techniques for Servers and Data Centers, *International Green Computing Conference* (IGCC), 2011

[76] D. Meisner and T. Wenisch. DreamWeaver: Architectural Support for Deep Sleep. *International Conference on Architectural Support for Programming Languages and Operating Systems* (ASPLOS), 2012

[77] B. Grot, D. Hardy, P. Lotfi-Kamran, B. Falsafi, C. Nicpoulos, and Y. Sazeides. Optimizing Data Center TCO with Scale-out Processors. *IEEE Micro*, Vol 32, Issue 5., 2012

[78] M. Guevara, B. Lubin, B.C. Lee. Navigating Heterogeneous Processors with Market Mechanisms. *IEEE International Symposium on High-Performance Computer Architecture* (HPCA), 2013.

[79] Y. Zhai, X. Zhang, S. Eranian, L. Tang, J. Mars. HaPPy: Hyperthread-aware Power Profiling Dynamically. *USENIX Annual Technical Conference* (ATC), 2014

[80] D. Lo, L. Cheng, R. Govindaraju, L. Barroso, C. Kozyrakis. Towards Energy Proportionality for Large-Scale Latency-Critical Workloads. *International Symposium on Computer Architecture* (ISCA), 2014

[81] A. Putnam, A. Caulfield, E. Chung, D. Chiou, K. Constantinides, J. Demme, H. Esmaeilzadeh, J. Fowers, G. Gopal, J. Gray, M. Haselman, S. Hauck, S. Heil, A. Hormati, J-Y. Kim, S. Lanka, J. Larus, E. Peterson, S. Pope, A. Smith, J. Thong, P. Xiao, and D. Burger. A Reconfigurable Fabric for Accelerating Large-Scale Datacenter Services. *International Symposium on Computer Architecture* (ISCA), 2014

[82] M. Shah, P. Ranganathan, J. Chang, N. Tolia, D. Roberts, T. Mudge. Data Dwarfs: Motivating a Coverage Set for Future Large Data Center Workloads. *Architectural Concerns in Large Datacenters Workshop* held in conjunction with the *37th International Symposium on Computer Architecture* (ISCA), 2010

[83] J. Chang, K. Lim, T. Mudge, P. Ranganathan, D. Roberts, M. Shah. A Limits Study of Benefits from Nanostore-Based Future Data-Centric System Architectures. *ACM International Conference on Computing Frontiers* (CF), 2012

[84] F. Chong. Bridging the Energy-Efficiency Gap in a Future of Massive Data. Keynote Address, *International Symposium on Performance Analysis of Systems and Software* (ISPASS) , 2014

[85] J. Bennet, H. Abbasi, P. Bremer, R. Grout, A. Gyulassy, T. Jin, S. Klasky, H. Kolla, M. Parashar, V. Pascucci, P. Pebay, D. hompson, H. Yu, F. Zhang, and J. Chen. Combining In-situ and In-transit Processing to Enable Extreme-Scale Scientific Analysis. *Supercomputing conference* (SC), 2012

[86] C. Trezzo, K. Webb, and K. Yochm. In-situ MapReduce for Log Processing, *USENIX Annual Technical Conference* (ATC), 2011

[87] H.Viswanathan, E. Lee, and D. Pompili . Real-time In-Situ Processing of Ubiquitous Mobile-Application Workflows. *IEEE International Conference on Mobile Ad-Hoc and Sensor Systems* (MASS), 2013

[88] 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

[89] P. Banerjee, C. Patel, C. Bash, A. Shah, M. Arlitt. Towards a Net-zero Data Center. *ACM Journal on Emerging Technologies in Computing Systems* (JETC), Vol. 8, Issue 4, 2012

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

[91] R. Singh, D. Irwin, P. Shenoy, K. Ramakrishnan. Yank: Enabling Green Data Centers to Pull the Plug. *USENIX Symposium on Networked System Design and Implementation* (NSDI), 2013

[92] X. Ma, P. Huang, X. Jin, P. Wang, S. Park, D. Shen, Y. Zhou, L. Saul, and G. Voelker. eDoctor: Automatically Diagnosing Abnormal Battery Drain Issues on Smartphones. *USENIX Symposium on Networked System Design and Implementation* (NSDI), 2013

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

[94] D. Wang, C. Ren, A. Sivasubramaniam, B. Urgaonkar, and H. Fathy. Energy Storage in Datacenters: What, Where, and How Much. *SIGMETRICS Performance Evaluation Review*, Vol. 40, No. 1, 2012

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

[96] B. Aksanli, T. Rosing, and E. Pettis. Distributed battery control for peak power shaving in datacenters, International Green Computing Conference (IGCC), 2013