Building Fuel Powered Supercomputing Data Center at Low Cost

Yiqing Hua[1,2] Chao Li[1] Weichao Tang[1,2] Li Jiang[2] Xiaoyao Liang[2]

[1]Sustainable Architectures and Integration Laboratory
Department of Computer Science and Engineering
Shanghai Jiao Tong University, Shanghai, China
E-mail: lichao@cs.sjtu.edu.cn

[2]Advanced Computer Architecture Laboratory
Department of Computer Science and Engineering
Shanghai Jiao Tong University, Shanghai, China
E-mail: liang-xy@cs.sjtu.edu.cn

ABSTRACT

Distributed power generations that fed with various economical clean fuels are emerging as promising power supplies for extreme-scale computing systems. Recent years have witnessed a growing adoption of these non-conventional power supplies in data center designs due to the heightening demand for reducing IT carbon footprint and server energy cost. However, the benefits of such a fuel powered data center are often severely compromised by its high initial capital cost (CapEx). This is because most pilot designs today either rely on expensive advanced generators or employ low-performance generators with costly standby power backup.

In this study we exploit heterogeneous generation to reduce the cost of data center powered by fuel. We show that different types of power supplies, if used together, can greatly improve the cost-effectiveness of self-generation but introduce a new layer of design complexity and raise an important question of how to dispatch computing tasks on heterogeneous power supplies. Specifically, due to the non-ideal output power response speed of heterogeneous generators, servers may incur serious power budget deficiencies when dispatching large amount of jobs. We refer to this phenomenon as *power lagging*, which jeopardizes system reliability and are not economical to be handled by costly power backup systems. To overcome this barrier, we propose *μBatch*, an agile load dispatching scheme that eliminate *power lagging* at the system/software level. Other than dispatch computing tasks in bulk without considering power system behaviors, *μBatch* intelligently splits job queue into small sets and incrementally schedule jobs based on the power ramping rate constraints and total power budget constraints. Using realistic HPC datacenter load traces, we demonstrate that *μBatch* enables supercomputers to smoothly operate on heterogeneous power. Our design helps data center operators save over 80% cost while maintaining the desired workload performance.

Categories and Subject Descriptors

C.0 **[Computer Systems Organization]**: General
C.4 **[Performance of Systems]**: Design Studies

General Terms

Design, Management, Performance, Efficiency

Keywords

Data Center, Dispatching, Hybrid Generation, Fuel, Cost

ICS'15, June 08–11, 2015, Newport Beach/Irvine, CA, USA.

DOI: http://dx.doi.org/10.1145/2751205.2751215

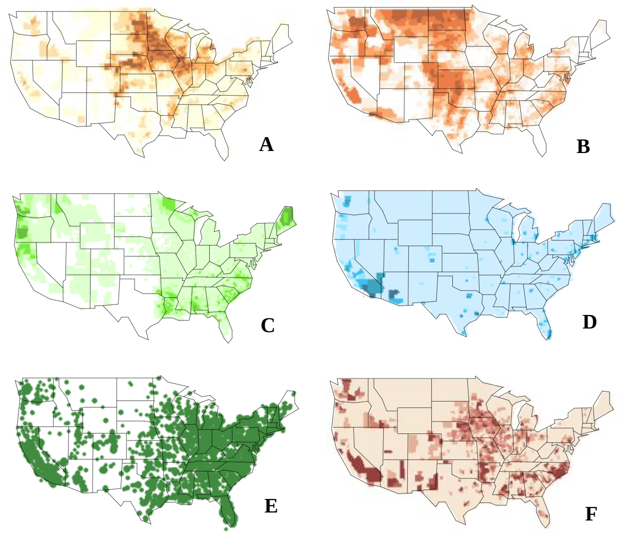


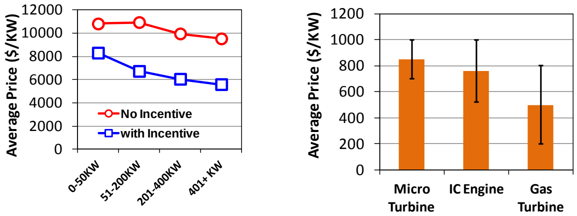
Figure 1. U.S. biomass resources maps. (A) Corn stover, (B) wheat straw, (C) forest residues, (D) wood/mil residues, (E) landfills, (F) manure. Data Source: National Renewable Energy Lab [4]

1. INTRODUCTION

With the advent of the smart grid initiative [1] and the evolution of various distributed generation (DG) systems [2], data centers today are facing new opportunities of supporting their escalating energy needs and continued capacity growth. Specifically, clean fuel based DG systems have shown great promise in offering 24×7 reliable power supply while maintaining carbon-neutral. If configured in combined heat and power mode (CHP) [3], these green generators can be over 80% efficient. Consequently, onsite fuel-based generation are expected to become the next frontier of power provisioning for server clusters and supercomputers.

In fact, there is a great potential to dramatically expand the use of clean fuel for power-hungry systems. Many fuels such as methane (natural gas), propane, ethanol, and methanol can be made from various biological materials (biomass) showing in Figure 1. The total biomass resource in the US is estimated to be 680 million dry tons per year, which can be exploited to generate about 730 billion kWh [5]. In 2010 the annual server energy demand in the US is about 135 billion kWh; in the worst case, this demand may double in 2015 [6]. Be that as it may, half of the biomass generated electricity is enough to power all the data center servers.

Many recent examples point to an increasing interest in data centers powered by various fuels. In Wyoming, Microsoft has deployed a non-conventional cluster for high-performance computing and modeling applications. Its power source is an onsite fuel cell system that converts methane biogas to electricity [7]. Several IT companies including Apple and eBay are also running part of their server racks on biogas-based fuel cell engine (called Energy Server



(a) Fuel cell costs (b) Heat engine costs

Figure 2. Capital cost of fuel-based generators [19-23]

[8]) developed by Bloom Energy [9, 10]. HP has investigated a data center power provisioning infrastructure that uses farm waste as primary fuel [11]. In addition, its Net-Zero data center can take advantage of various bio-gas turbines as well [12].

Given the growth in the popularity and scale of clean fuel based DG systems, this study aims to develop an understanding of fuel-powered large-scale computing facilities. We do not argue that tapping into fuel is the only design choice or the best way of delivering green high performance computing. However, the direct use of clean fuel in data center is attractive for several reasons:

1) The energy cost is usually low, particular for CHP systems (a.k.a. cogeneration) that have higher efficiency [3]. There is a favorable differential between the price of natural gas and the price of electricity, called the spark spread [13]. As a result, the annual energy cost savings can be up to $700 per kW installed capacity [3].
2) The reliability of fuel delivery system is known to be high [14]. Unlike the utility power grid, the gas system is underground and exposed to fewer natural disasters. Its delivery contracts often exhibit over 99.99%, which matches the availability of a fully fault-tolerant Tier-IV data center and is much better than the main grid (99.9% or less) [15-17].
3) Onsite fuel based generators are often dispatchable, namely, their power output can be adjusted on demand [3]. Each generation module can also be turned on or off dynamically based on the average loading of data centers. This is a key advantage over wind/solar power systems which have intermittent and time-varying output.
4) Supporting smart energy initiatives can be a wise business move as well. Given the growing awareness of climate change and energy security, companies and consumers will increasingly show preferences to a diverse generation method and smarter grid [18]. Integrating various carbon-natural fuels into a datacenter facility's structure helps attain leading market positions and strong public relations.

One of the key obstacles that could hold back the development of fuel-powered data center is high cost. Existing fuel-based DG system can be grouped into two broad categories: fuel cells (*FC*) and heat engines (*HE*). The cost of a fuel cell system is inherently high, as shown in Figure 2-a. Without incentive, a 100 kW system sold by Bloom Energy can cost $800,000 [24]. Although heat engines are about one-tenth the price of fuel cells (Figure 2-b), they have slow output response speed and often require costly standby power capacity. For example, if using utility power grid and redundant UPS capacity as backup, the power infrastructure cost is estimated to be $10~25 per watt [25]. In addition, utility companies also charge a power demand fee from $8/kW [26] to 18/kW [27] per billing cycle. Moreover, if not managed properly, UPS batteries incur frequent replacement cost at over $2,000 per kW [28]. Therefore, many potential benefits of the fuel-powered data center could be severely compromised due to high power related cost.

To build economically viable and even cost-competitive system, this paper explores mixing different power generation technologies and minimizing underutilized standby power in fuel-powered data centers. We propose a new power provisioning strategy called Phi (Φ), which stands for *provisioning in hybrid and islanded mode*. This architecture has two salient features. (1) It exploits a hybrid power infrastructure that combines fuel cells with inexpensive heat engines. The resulting system is expect to exhibit a synergism in which the data center can yield a cost-effectiveness far better than that can be offered by either power generator alone. (2) It takes advantage of the highly reliable nature of the DG systems and offloads the burden of managing power spikes onto them. This could eventually allow us to eliminate the needs of reserving large power capacity from the utility grid and onsite battery (i.e., islanded).

While the Φ architecture has great promise for reducing the CapEx of power infrastructure, applying it to a mission critical system can be challenging. The main reason is that hybrid power systems, from the viewpoint of a sensitive load, are not ideal power supplies. A homogeneous fuel cell system only take several seconds to ramp power up as load changes. Therefore, it is reasonable to completely replace backup power with fuel cell [17]. Nevertheless, most heat generators takes several minutes or more to ramp up since the change of the turbine speed takes time. With majority being inexpensive heat engines, the Φ architecture has very limited capability to handle unexpected load surges. Without careful control, existing data centers can get very sluggish power increase when schedule large amount of tasks in a short time, which we refer to as "power lagging". This phenomenon can often cause suboptimal energy efficiency and even costly power failures.

To support a graceful transition to the Φ architecture, we have developed a novel load dispatching scheme called *μBatch*. It is an agile load dispatching scheme tailored to the behaviors of hybrid power supplies. Rather than dispatch batch jobs merely based on the availability of computing resources, *μBatch* also considers various supply-side constraints. It firstly divides a large job queue into two medium job sets based on the job classification information and the available power budget of the two types of DG system (e.g., fuel cell and heat engine). It then incrementally schedule each job set based on the power ramping rate constraint of the power supply. This judicious load dispatching scheme enables fuel-powered data center to overcome the "power lagging" barrier while maintaining high performance and efficiency.

To our knowledge, this is the first extensive study of a fuel-powered supercomputing data center design at the architecture and system level. This paper makes the following contributions:

- We investigate clean fuel powered high performance data centers. We propose Phi (Φ), a new power provisioning architecture which allows a data center to maximally exploit the benefits of fuel-based generation at low cost.
- We identify the "power lagging" issue in data centers powered by hybrid DG generators. We characterize it and show that being able to adapt to the "power lagging" is the key to build a cost-effective system.
- We propose μBatch, an agile load management framework tailored for fuel-powered data centers. It is a power-driven, incremental job dispatching approach that allows one to overcome the limitations of power supply through a software and system level control.
- We evaluate the design space of μBatch using realistic job traces. We show that it can reduce over 80% TCO in a fuel-powered data center, with near-optimal (within 1% difference) performance.

The rest of this paper is organized as follows. Section 2 discusses related work. Section 3 introduces background of fuel based generation. Section 4 propose and analyze the Phi architecture. Section 5 elaborates the μBatch load dispatching scheme. Section 6 describes experimental methodology. Section 7 presents evaluation results. Finally, Section 8 concludes this paper.

2. RELATED WORK

[Data Center Powered by Fuel] Very little prior work exists on fuel-powered data center design. Similar studies are mainly carried out by several pioneers in the industry. For example, Sharma et al. [11] from HP investigate data center self-generation that uses biogas generated from livestock wastes. Their work introduces the key resource management flow but lacks quantitative analysis. [17] conducted by Microsoft is the first research study to promote fuel based data center design. Zhao et al. [29] also from Microsoft have demonstrated the feasibility of operating server racks on fuel cells. These designs choose to conservatively add battery as power backup. In contrast, our design allows the data center to further reduce the high capital cost by minimizing the needs of backup power capacity. In general, all these prior work focus on the supply-side infrastructure and power management from an electrical point of view. In this paper we propose workload-side solutions that enable us to maximally exploit fuel-powered data center.

[Load-Following Power Supply] Another representative work in the context of clean fuel powered data center is to explore the use of the load following capability of generators. For example, Liu et al. [30] formulate an energy-aware request routing problem in geographically distributed cloud data centers powered by fuel cells. Li et al. [31] exploit distributed generation for tracking the fluctuation of data center demand. Unfortunately, both designs only look at homogeneous power provisioning schemes and miss the opportunity to reduce high design cost. In addition, our power management scheme considers the power ramping rate constraints of load-following DG, while prior optimization schemes in these work overlooked this critical power supply characteristics.

[Non-Dispatchable DG System] There is a large body of work in terms of intermittent (non-dispatchable) renewable energy (solar and wind energy) powered data centers [32-38]. Sharma et al. [32] investigate the impact of power supply intermittency on the most critical memcached workloads. Goiri et al. [33] propose to leverage deferrable batch workload in HPC centers to maximally harvest the intermittent green energy. Li et al. [34], on the other hand, uses VM-based workload migration to improve the utilization of wind energy powered server cluster. Deng et al. [35] show that the profit of a green data center can be largely affected by the way they manage carbon footprint. Gao et al. [36] propose FORTE, a framework that can determine suitable expansion plan. A data center expansion strategy is also proposed in [37], which leverages modular solar panels and distributed battery systems. In [38], the author has demonstrated a scaled-down prototype of a grid-connected solar data center called Parasol. Our work differs from these studies in that we focus on another important green power supply: clean fuel-based dispatchable generators. In addition, we investigate the most effective way to cut capital cost in these systems other than merely emphasize improving green energy utilization or performance.

[Hybrid Power Provisioning] A few recent studies have focused on hybrid power supplies. For example, Li et al. [39] investigate data centers on green energy mix that combines dispatchable and non-dispatchable power. However, they still focus on homogeneous fuel-based generators and rely on bulk backup batteries. In [40], Liu et al. propose heterogeneous power provisioning. Their work is only limited to hybrid energy storage devices (i.e., batteries and super-capacitors). In other domains, fuel cell/heat engine hybrid systems have been well studied with the aim to improve generator inner efficiency [41, 42]. In this study we show that simply connecting computing loads to these systems often result in sub-optimal design tradeoffs due to a lack of supply-load coordination. When hybrid power system become part of a data center's energy portfolio, new power-aware load management schemes at the computer architecture and system level is highly desired.

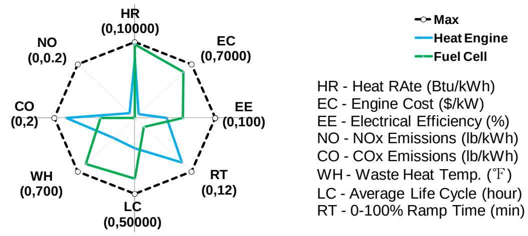


Figure 3. Comparison of fuel-based generation. Each technology has its own pros and cons [19-23, 46-48]

[Batteries and Peak Shaving] Prior research has demonstrated the benefits of leveraging energy storage devices, such as UPS batteries, in datacenters to shave peak power demand [43-45]. This can significantly save OpEx and Cap-Ex cost. However, batteries do not generate power on their own and they can only serve as ancillary backup in distributed generation systems. To minimize the upfront cost in a fuel-powered data center, our goal is to minimize the dependency on them by leveraging smart load dispatching.

The novelty of our work is three-folded. (1) We explore powering high-performance computing data center with hybrid clean fuel based generation. (2) We follow a KISS[1] design principle when designing the Phi (Φ) architecture – minimizing unnecessary power backup makes our system more cost-effective. (3) We propose a new load dispatching scheme called μBatch to provide a system support for managing heterogeneous power supply.

3. BACKGROUND

This section briefly introduces fuel-based self-generation which has the potential to reshape the way we run data centers.

3.1 Generating Power from Fuel

Electricity can be generated from fuel in different ways. Fuel cells produce electricity through a pollution-free electrochemical process [42]. Heat engines, one the other hand, rely on various advanced combustion turbines that burn fuels [46]. Both fuel cells and heat engines generate large amount of heat, which can be recycled to yield an overall energy efficiency of over 80% [47]. Consequently, they are deemed to be among the most promising distributed generation resources in the smart grid era [47].

No generation technology totally dominates. The radar chart in Figure 3 compares several key parameters of typical fuel cells and heat engines. Heat engines offer lower acquisition cost while fuel cells have much better power ramp rate (the speed to change power output). Sudden and massive requests for computing nodes can cause significant power demand increase, which challenges the ramp rate of distributed generation systems. If this demand-supply

[1] KISS: Keep It Simple, Stupid. *http://en.wikipedia.org/wiki/KISS_principle*

Generation Technology		Warm-up Time (cold start)	Ramp Rate (0-100%)
FC	Solid Oxide FC	10 min	sever seconds
	Phosphoric Acid FC	6 min	sever seconds
	Proton Exchange Membrane FC	1 min	sever seconds
HE	Micro Turbine	7 min	8 min
	Gas Turbine	7 min	8 min
	Internal Combustion Engine	1 min	6 min

Table 1. Power output response speed comparison

power gap is not well managed, the result can be significant deterioration in power quality or even costly power outages. In Table 1 we show the ramp rate of several representative technologies [46, 47]. Fuel cells can release power almost immediately and could be used to handle abrupt power changes once they are warmed up. In contrast, most heat engines need time to be committed (reach a desired engine speed). They provide a sluggish load following, which typically occurs on the timescale of several minutes.

3.2 Advantages and Limitations

The technical and economic viability of tapping into fuel largely depend on the availability of fuel supplies. Fortunately, there are a variety widely distributed biomass resources such as agricultural residuals, forests residuals, industry byproducts, and urban wastes. In addition, the extensive natural gas infrastructure also provides very convenient low-emission fuel delivery.

Clean fuel-based onsite generators are ideally suited for high-performance computing data centers for several reasons. (1) They are reliable source of power. A fuel cell generator can be expected to run continuously for over ten years [48]. In fact, the ability of an aging electric power grid to meet the soaring server power demand is at stake [49, 50]. Data center operators are increasingly challenged to rely on conventional approaches to meet their expanding power capacity needs. (2) Fuel generated from biomass is renewable energy resources which greatly reduce the carbon footprint of a supercomputing center. Although burning biofuel releases CO_2, this is offset by the CO_2 absorbed in the growth of new biomass through photosynthesis. (3) They can provide dispatchable power whenever data center have different power demand. Not like intermittent solar/wind energy, fuel-based generators produce power in a controllable fashion. Various onsite generators can be interconnected to form a low-voltage network called micro-grid [47], which further increases the manageability of onsite generation.

The major barrier to the adoption of large-scale fuel-based generation in data centers is high initial cost. In some scenarios, renewable energy certificates (carbon credits) could turn it feasible to generate energy onsite using clean fuel. Still, the price of fuel cell and the price of backup power system that is required by heat engines often make onsite generation economically less attractive. Substantial effort has been put into optimizing the inner efficiency of the generator. However, improving cost-effectiveness in fuel powered data center demands continued innovation in architecture design and system management. Our work fills this critical void.

4. THE PHI (Φ) ARCHITECTURE

To turn fuel-powered HPC data center into a competitive edge, we explore aggressively utilizing onsite generation while minimizing the needs of power backup. We propose a new power provisioning architecture Phi: power provisioning in hybrid (heterogeneous) and islanded (isolated) mode. The Φ architecture allows us to greatly reduce the cost of a fuel-powered HPC data center.

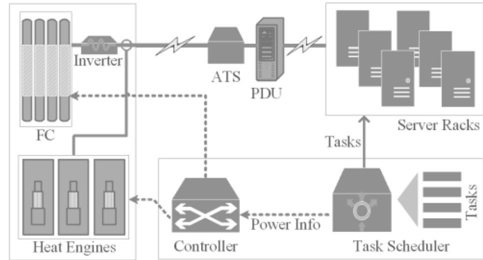


Figure 4. Schematics of the Phi architecture

In this section we first introduce the unique architecture of Phi. We than discuss its necessity and cost benefits in data center design. Finally we discuss the main limitation of this architecture from a power management point of view. We propose a system-level solution to overcome this limitation in Section 5.

4.1 Hybrid and Islanded System

The main feature of Phi is two folded. First, Phi employs heterogeneous power. Second, Phi is isolated from other standby power sources (no matter it is utility power grid or onsite batteries).

Figure 4 depicts the overall structure of Phi. The hybrid power system is the only power source that the data center is connected to. The output of fuel cells is DC power. It is first inverted to AC and then interconnected with the heat engine to form a local micro-grid system. On the data center side, a power controller is responsible for coordinating the output levels of different power supplies.

Combing fuel cells and heat engines brings unprecedented opportunity for data center power provisioning. Such a heterogeneous power system can optimize global efficiency and cost-effectiveness. The total capacity of the installed fuel cell is normally larger than that of heat engines. Fuel cells have better power ramping capability and can be used to handle the abrupt power demand change. Heat engines are inexpensive source of power and are used to provide base load power and a slower load following service.

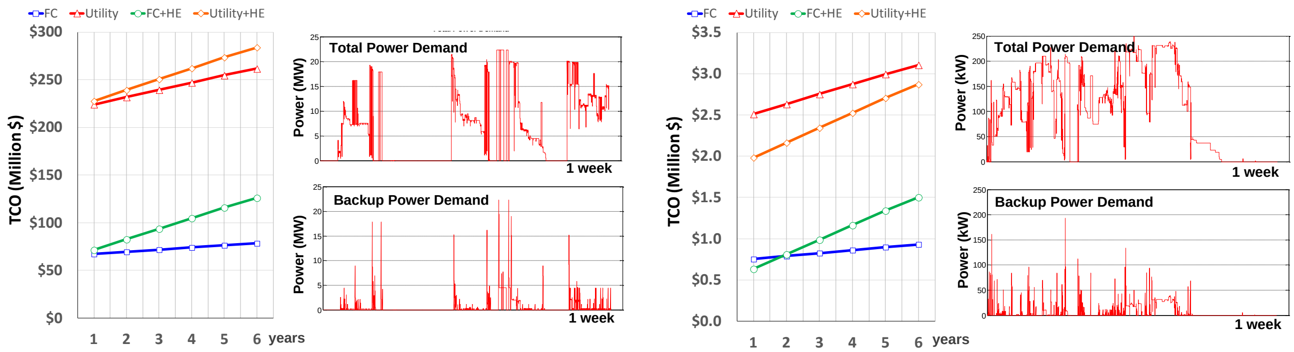
The islanded mode of Φ is partially enlightened by the islanding state of a micro-grid. A micro-grid can disconnect itself from the main grid and enters an islanding state when the utility grid incurs power quality problems. It maintains normal operation on all of its loads in an islanding mode. In this study our goal is to enable data center to run smoothly on islanded mode to radically reduce cost.

It is not uncommon to use standalone (i.e., islanded) systems. For example, Yank [51] allows a green data center to unplug from the utility grid. Besides, iSwitch [34] also has partial of its server clusters powered only by onsite wind turbines. However, both designs have backup computing capacity. They move workloads to reserved computing systems when renewable power is inadequate or power failure happens. In this study we do not over-provisioning computing systems (with is very costly) and therefore do not have additional server racks as backup. In fact, there is no need to do so on fuel powered data center since the availability of energy supply is high (99.99%) and the power supply is stable.

4.2 The Power Lagging Problem

Although the Φ Architecture is expected to greatly reduce the design cost, current data centers are not able to benefit from it. The reason is very simple: they still need costly backup power system in order to run smoothly on the Φ Architecture.

Sluggish power ramp speed of the hybrid power system is the main challenge in directly employing Φ in data centers. Heat engines are the majority in the Φ Architecture and they have very long ramp up time. Fuel cells can increase its output in a matter of



(a) Blue Gene/P system called Intrepid (b) San Diego Supercomputing Center

Figure 5. TCO analysis of different power provisioning schemes

seconds to adapt to rapid shocks in power demand, but their total capacity is very limited. In other words, the Φ Architecture itself simply lack the capability to many data center power surges.

Consequently, data centers encounter a situation in which the power supply lags behind the power demand. Existing data center load dispatchers are unaware of the power ramping behaviors of the onsite generation, and therefore release large amount of jobs upon the availability of computing resources. This can introduce a step change in data center power demand. Oftentimes, the huge power mismatch between demand and supply cannot be handled by the Φ Architecture either due to inadequate fuel cell capacity or slow power ramp rate of the heat engine. This situation is often worsened by under-provisioning fuel cells.

4.3 Understanding the Cost Overhead

Due to the power lagging issue, the cost-effectiveness can be still very low in data centers employing Φ. To understand this, we characterize the total cost of ownership (TCO) of various system configurations, as shown in Table 2. The capital cost of fuel cell is conservatively estimated to be $3000/kW, and its O&M cost is at $0.04/kWh [52]. The infrastructure cost of the utility backup infrastructure is $10/W and the electricity price is $0.13/kWh. Although the engine cost of heat engine ($200/kW) is much lower than a fuel cell, the O&M (operating and maintenance) cost can be higher if using conventional combustion engine set ($0.2/kWh) [53].

Strategy	Descriptions
UG	Only use the utility power grid
FC	Only use FC as power supply
HE+UG	HE with utility grid as backup
HE+FC	Φ without optimization

Table 2. Evaluated design scenarios

We use realistic HPC data center traces obtained online [54]. One trace comes from a BlueGene/L system at Lawrence Livermore National Lab. It features a massive amount of short jobs (over 90% jobs run for less than 5 seconds), resulting in heavily fluctuating power demand. The other trace is collected from the San Diego Supercomputer Center. Its average job size (requested number of CPU nodes) is about 1/50 of the BlueGene/L system, but its average job runtime is longer (almost 2X). We provision each data center in such a way that the data center power system can just satisfy the server power demand without affecting the job completion deadline. Figure 5 presents our evaluation results.

It is clear that a fuel-powered data center can exert heavy power spikes on its backup power systems. As shown in Figures 5-(a) and 5-(b), both HPC data centers can create highly bursty total power demand. Since heat engines have very limited ramp up rate, *HE+UG* has to occasionally request large amount of power for a very short period (several minutes) from the utility grid. While the total energy drawn from the utility grid in this case is negligible, the associated infrastructure cost can be skyrocketing. In addition, the utility company often installs a power demand meter onsite. A demand charge will be applied based on the highest average kW measured in a 15-minute time interval in the last billing period [26]. Therefore, the usage of heat engine as a primary power supply in the *HE+UG* does not reduce its reliance on utility power backup. Consequently, the TCO of *HE+UG* can be even higher than a conventional data center that is only powered by the utility grid.

Similarly, the TCO of the *HE+FC* configuration is also not optimistic at all if data center operators overlook the power lagging issue. As Figure 5 shows, the cost of *HE+FC* may significantly exceed a fuel-cell-only data center in a typical five-year *HE* life span. This is because server loads can create significant amount of power surges, which requires large amount of onsite fuel cells to provide transient power support. Thus, the usage of heat engine as a primary power supply in the *HE+FC* configuration does not reduce a data center's reliance on fuel cell system. On the country, the result is significantly over-provisioned onsite generators.

In addition, simply adding battery systems to the *HE+FC* configuration is not a sound solution. First, batteries can be very costly if used as transient power backup; their power capacity cost can be $2000~4600/kW for the most widely used lead-acid batteries [28]. In addition, frequent high-current discharge greatly reduces a battery's lifetime, which further increases the depreciation cost of a battery. Moreover, batteries are essentially not power generators. They need to be charged once their stored energy level is low and they also incur a round-trip energy loss issue.

In fact, a very easy and intuitive way to reduce TCO in data centers running on Φ is to reduce the height of any abrupt power change. In Figures 5-(a) the height of power spikes exerted on the backup system is almost the same as the maximum power demand of the data center. In Figures 5-(b), however, the height of the peak backup power demand is only 87% of the peak data center power. As a result, the *HE+UG* configuration in Figure 5-(b) shows better cost effectiveness. The *HE+FC* configuration in Figure 5-(b) also becomes cheaper compared to *FC* in the first two years.

Summary: Current data centers cannot directly run on the Φ architecture due to a lack of resilience to the power lagging problem of hybrid generators. They often requires significant amount of backup power capacity, which contradicts our commitment to reduce the cost. The cost of a fuel-powered data center can be greatly affected by the way we manage the data center load.

5. INCREMENTAL LOAD DISPATCHING

We investigate how load management will help to improve a data center's resilience to the power uncertainty incurred in Φ. Rather than waiting for the power systems to be smart (i.e., better *HE* and cheaper *FC*), we instead design an intelligent load dispatching strategy. It allows data centers to develop the flexibility to run smoothly on Φ, in spite of the fact that the power supply is less-than-perfect and the power demand unavoidably fluctuating.

5.1 Overview of µBatch

We propose *µBatch*, an agile load dispatching scheme tailored to the behaviors of hybrid distributed generation. Traditional dispatchable power systems have been designed in this way that generation can follow the load power demand. To enable a data center to graceful operate on Phi, it is best to also dynamically control the job flow so that the load also follows the generation patterns to some extent. To achieve this, µBatch uses an incremental job dispatching strategy. It judiciously dispatch jobs based on the power ramp rate and power budget constraints.

Existing data center job dispatching strategies are largely workload-driven. A list of jobs are normally submitted to a scheduler that is responsible for validating job parameters, queueing jobs, allocating compute nodes, and monitoring the status of job operation. One important goal of conventional scheduling policies is to reduce ensure better resource utilization while minimizing job turnaround time. For example, backfilling policies [55] rearrange jobs instead of execute jobs in their order of arrival to minimize system fragmentation. Such a compute load oriented job scheduler typically dispatches jobs once the computational resources are available in the system. They can handle complex parallel jobs with heavily constrained computing resource but cannot efficiently adapt to the power lagging issue in hybrid power provisioning environment.

In contrast, µBatch is primarily power-driven. It takes into account the available power capacity of different types of power supplies (i.e., heat engine and fuel cell) and their power ramping speed. It then classifies the submitted jobs and evaluates all the computational resources available in the data center to decide which jobs have to be dispatched, where and when. An important goal of µBatch is to optimally exploit the available energy source while minimizing the negative impact of power lagging on job execution. Its design follows three general considerations from the view point of data center power management:

First, we do not over-generate power and energy, since electricity cannot be stored onsite for Phi. In fact, over-generating is not wise even for grid-connected data center. Although the data center owner could run *FC* and *HE* at full capacity all the time and export excess power to the utility company, the price the utility pays for the exported power does not make it economically attractive to do so. Therefore, it is best to full utilize all the energy generated.

Second, we defer workload, other than curtail workload. In other words, µBatch schedule jobs only when the power generation ramps up to a desired level. Similar to existing work [33], we take advantage of the fact that most data processing jobs are deferrable to a certain extent. Existing designs often resolve power brownout (caused by over-loading) in a rather crude way: using DVFS to place a hard power cap, or simply shut down the server (load shedding). We chose not to do so since running at less-than-ideal frequency significantly decrease the efficiency of a server. In addition, scaling node performance or shutting down servers can cause great adverse impact on all the parallel tasks that have interdependence.

Third, we manage the power usage of different types of generators at the system level. There has been prior work that divides data

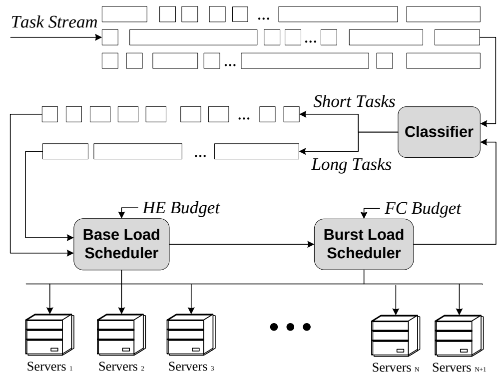


Figure 6. Schematics of the µBatch dispatching strategy

center into two server clusters and uses separated power provisioning paths [34]. In fact, such a hardware-based approach is not suitable for fuel-based generation since it may exacerbate the power lagging issue on the sever cluster that connect to heat engines. Workload migration may be a solution to allow power sharing between two clusters, but it incurs significant overhead.

5.2 Dispatching Strategy

We formulate the job dispatching strategy on Φ as a two-fold problem. 1) Given limited power ramp speed, apply appropriate load deferring policy for jobs that have different runtime. 2) Given that the installed fuel cell capacity is very limited, appropriate power capacity management between responsive (i.e., fuel cells) and sluggish (i.e., heat engines) power supply is necessary.

Figure 6 illustrates the proposed µBatch design. Its scheduling approach takes three steps: classification of forthcoming job requests, allocation of base-load power (i.e., heat energy power capacity) for majority of the jobs, and opportunistic usage of fuel cell power capacity to manage any potential bursty loads. Thus, µBatch can be seen as a supply-load co-scheduling technique.

In the first step, µBatch requests profiling information from the job queue and classifies jobs into two types: short jobs that are expected to finish within the lead-in time (i.e., power ramp up time) of HE generators, and long jobs that maintain steady power demand for extended period of time (> than the lead-in time). Short jobs are more likely to result in transient load power demand fluctuation, while long jobs contribute less to load fluctuation.

Our dispatcher first allocates the heat engine generation. Heat engines accounts for a larger portion of the available power capacity and are inexpensive source of power. They are ideal base-load power for long-running jobs. In Figure 6, the base-load scheduler dynamically checks the newly available power budget as a result of power ramping. Once the additional power budget headroom is meet the power demands of any one job in the queue, our scheduler will dispatch it immediately no matter it is short or long.

Our dispatcher allocates the fuel cell generation after the heat engine power has been allocated. Fuel cell power capacity is limited and it is critical to judiciously utilize this precious source of responsive power supply. In Figure 6, our bursty-load scheduler gives high priority to short jobs high priority. It selects the most delayed one from short jobs that can be scheduled based on the available fuel cell power budget headroom. Instead, we never use fuel cell to power long jobs for good reasons: the impact of job deferring on long jobs is much smaller than short jobs. Since most heat engine can ramp up to full capacity in 10 minutes, the maximum job delay for a long job can be 10 minutes which may be a

Algorithm 1: Pseudo code for μBatch dispatching

```
Indicator = Maintain;
for each fine-grained timestamp
    change the power supply based on the Indicator;
    add the job to wait[];
while HE has enough power
    schedule a job from wait[];
if wait[] is not empty
    Indicator = Up;
else if  Power goes down won't affect job running now
    Indicator = Down;
else Indicator = Maintain;
while FC has enough power
    schedule the most delayed short task  from wait[];
```

Trace Abbr.	Short/Long	Highest Peak	
Seth	24%	1573KW	Small
BH	68%	1768KW	Small
CM5	61%	3415KW	Medium
PIK	20%	5536KW	Large
DS	52%	7124KW	Large

Table 3. Summary statistics of the selected traces

very small percentage of its overall all execution time (several hours or many days). In contrast, a 10-minutes delay may be too long for jobs that are expected to finish within tens of seconds.

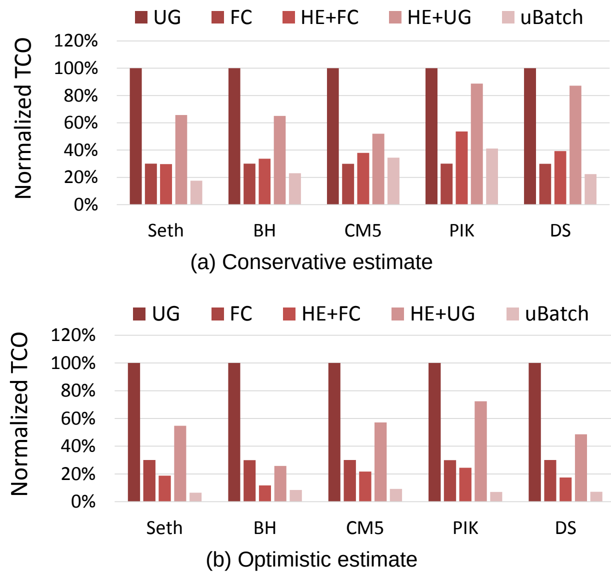
In Algorithm 1 we show the pseudo code for the co-scheduling scheme used in μBatch. Similar to existing backfilling policy, we expect users to provide nearly accurate estimates of their job execution durations. In each round of scheduling there are two stages: base-load allocation and bursty-load allocation. Jobs that are not dispatched will remain in the job queue. The result is a series of micro/small batch of jobs and an incremental job execution flow.

We expect μBatch to reshape data center power demand. It maintains a mild power demand variation, and therefore enables data centers to run smoothly on Phi, our clean fuel-based, sustainable power provisioning architecture with very low cost.

6. EVALUATION METHODOLOGY

We use data logs from six large-scale clusters [54]. Each log provides the job arrival time, waiting time, the number of processors requested, the actual duration, average CPU time, and the start time. Our framework uses discrete-event simulation approach to process a chronological sequence of job events. For each trace we evaluate 5 million seconds of operation duration, which is about two months. We select six traces as summarized below:

- **Seth**: *This log is from a 120-node Linux cluster named Seth, which belongs to the HPC Center North (HPC2N) located in Sweden. Original job scheduler is Maui*
- **BH**: *This is the San Diego Supercomputer Center Blue Horizon log. The total machine size is 144 nodes (8-way SMP) for batch use. Jobs are collected using the LoadLeveler.*
- **CM5:** *This trace comes from a 1024-node connection machine CM-5 system deployed at the Los Alamos National Lab. Scheduling was performed by the DJM software.*
- **PIK**: *This log is collected from a 320-nodes IBM iDataPlex cluster at the Potsdam Institute for Climate Impact Research (PIK). The maximal job size observed used 128 nodes.*
- **DS**: *The San Diego Supercomputer Center DataStar log contains data comes from a 184-node machine. However only 171 nodes are used to process batch jobs.*



(a) Conservative estimate

(b) Optimistic estimate

Figure 7. Normalized TCO schemes for different schemes

Two parameters affect our evaluation: 1) the ratio between short and long jobs, and 2) the highest power demand of the system. The former is closely related with the scheduling decisions of μBatch, while the latter significantly affect the total power capacity planning decision. Table 3 summarizes the statistics of the selected traces. We classify jobs according to their execution duration. The threshold that separates short jobs and long jobs is 10 minutes. Significant variability is observed in the job duration. The ratio between short and long jobs ranges from 7% to 68%. We consider both small-scale clusters and extremely large data centers.

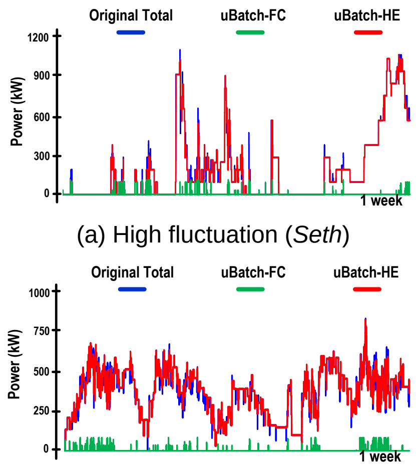
The data center infrastructure we evaluated is based on the HP ProLiant DL 380 G6 server system (Intel Xeon L5530 processor, 2.4G). Our simulation assumes that the jobs are CPU-intensive. According to the published SPECpower_ssj2008 Results, the evaluated server system has an average power of 93 Watt per chip at 100% utilization and a minimum power of 31 Watt per chip in active idle state. The total onsite power capacity of the heat engine system is based the maximum peak power of the evaluated data center trace. We assume the worst-case ramp time is 10 minutes. The default fuel cell capacity is one-tenth of the heat engine.

7. RESULTS

In this section we discuss the benefits of applying μBatch to data centers powered by hybrid onsite green power supplies. We compare μBatch to various design approaches discussed in Table 2. We first evaluate the cost benefits of the proposed Phi architecture. We then discuss the performance impact of μBatch. Finally we evaluate the importance of capacity planning in our design.

7.1 Cost Benefits

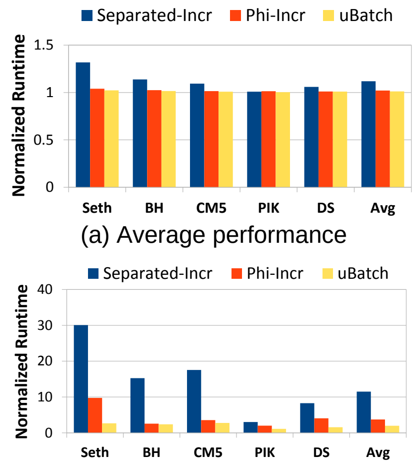
The key motivation of the Phi architecture and μBatch is to reduce the cost of a fuel-powered data center by enabling aggressive utilization of the power provisioning systems. In Figures 7-(a) and 7-(b) we present the 5-year total cost of ownership for different data center traces under various power management schemes. We consider both conservative design estimation and optimistic design estimation (i.e., the O&M cost of heat engine is $0.2/kWh [53] and



(a) High fluctuation (*Seth*)

(b) Mild fluctuation (*CM5*)

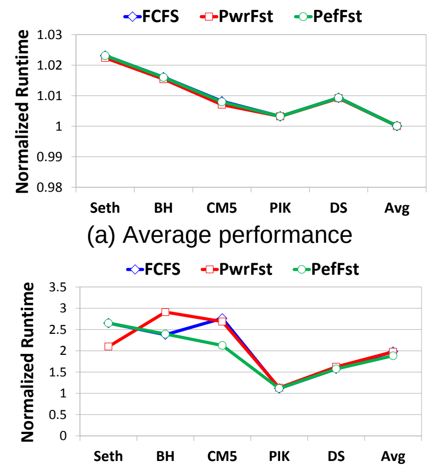
Figure 8. Power re-shaping results



(a) Average performance

(b) Worst-case observed

Figure 9. Workload performance



(a) Average performance

(b) Worst-case observed

Figure 10. Impact of scheduling policy

\$0.03/kWh [56], respectively). In Figure 7-(a), µBatch shows much lower cost compared to *HE+FC*, but in some cases it shows higher than a fuel-cell-only design. For CM5 and PIK in Figure 7-(a), the initial cost of µBatch is only 19% and 26% of *FC*, respectively. However, their TCO values under µBatch become larger than *FC* in the fourth year of deployment. Therefore, if the heat engine has high O&M cost (e.g., when using bio-diesel generators of low lifetime), the benefits of hybrid power system can be significantly compromised.

When factor in the low price of some heat engines, µBatch could significantly outperform other designs. In Figure 7-(b) we consider heat engines (e.g., gas turbines) that are more durable than conventional combustion engines and not sensitive to fuel quality like fuel cells [11]. If data centers can use methane biogas from a landfill or waste treatment plant (like Microsoft), the fuel cost can be even lower. In fact, even a small drop in energy cost can deliver substantial cost savings over time. On average, the TCO of the proposed fuel powered data center (i.e., Phi + µBatch) is only 7% of constructing a new utility-based data center. Its cost is only 30% of a fuel-cell-only data center (e.g., data centers running on the Bloom's Energy Server) and 19% of an non-optimized hybrid power provisioning approaches (i.e., *HE+FC*)

In Figure 7, the cost improvement of µBatch varies for different data center traces. For example, the normalize cost of *Seth* is only half of *CM5*. This is primarily becomes these data center traces have different power demand behaviors. Figures 8-(a) and 8-(b) show the power demand of the original data center traces and the actual power provisioning schemes under µBatch. It is clear that Seth has many significant, abrupt power demand changes, while CM5 has relatively smaller power variation range. As a result, the optimization effectiveness of µBatch on Seth is more noticeable.

7.2 Performance

The design of Phi and µBatch is not at the cost of workload performance. To demonstrate this, we further compare µBatch with two other important design alternatives: (1) *Separated-Incr* applies the proposed incremental load dispatching on a conventional separated power provisioning scheme similar to [34]. (2) *Phi-Incr* uses incremental load dispatching on the proposed Phi power provisioning architecture, but without the 3-step optimization of µBatch (i.e., job classification, first-round base-load scheduling, and the second-round bursty-load scheduling, detailed in Section 5.2).

Our results show that µBatch provides better performance guarantee. In Figure 9-(a) we show the average normalized job turnaround time (normalized to its original duration). On average, *Separated-Incr* yields a job turn-around time increase of 12%, while the number for *Phi-Incr* is just above 2%. In contrast, µBatch shows only 1% compared to the original job turn-around time. In fact, the most significant benefit of µBatch is to provide a better SLA even in the worst-case. The performance differences among the worst delayed job are much larger, as shown in Figure 9-(b). *Separated-Incr* could cause up to 30X increased duration on some short jobs that requires only tens of seconds to finish. In contrast, in the worst case µBatch maintains less than 10X job turn-around time increase, on average this number is less than 1X (98%).

To further under the benefits of µBatch and the impact of various job scheduling policy used in µBatch, we evaluate three variations of µBatch: *FCFS*, *PwrFst*, and *PefFst*. They affect the way we schedule jobs in the base-load scheduling and bursty-load scheduling steps (detailed in Section 5.2):

- ***FCFS****: jobs are always selected in a first-come-first-service way if multiple jobs can be scheduled based on the available power headroom*
- ***PwrFst****: jobs that have less power demand are given higher priority if multiple jobs can be scheduled based on the available power headroom*
- ***PefFst****: jobs that have the longest waiting time are selected first if multiple jobs can be scheduled based on the available power headroom. It is the default policy for µBatch*

In Figures 10-(a) and 10-(b), while the differences of the average job turn-around time are not noticeable, the differences in the worst case we observed can be large. *PefFst* always show better worst-case job turn-around time than *FCFS* (6% improvement on average). However, in some cases *PefFst* outperforms *PwrFst* while in other cases it does not (such as Seth). This is because *PefFst* may select a job that has very high power demand, and Seth happens to have many such short-and-large jobs, as indicated by Figure 8-(a). This could prevent other delayed jobs from being timely dispatched. Therefore, if users are sensitive about the worst-case performance, one might need to design a more complicated policy that considers both job power demand and waiting time. Otherwise, either *PwrFst* or *PefFst* can provide satisfactory performance guarantee (on the average job turn-around time)

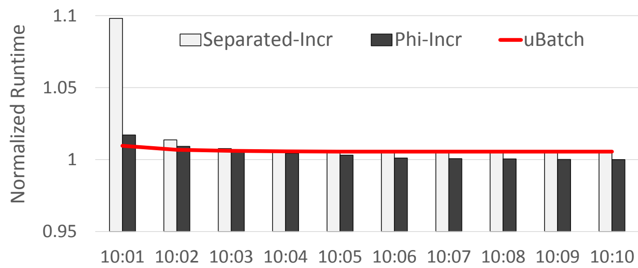


Figure 11. Average perfomance under different capacity ratio between *HE* and *FC* (the capacity of *HE* is fixed)

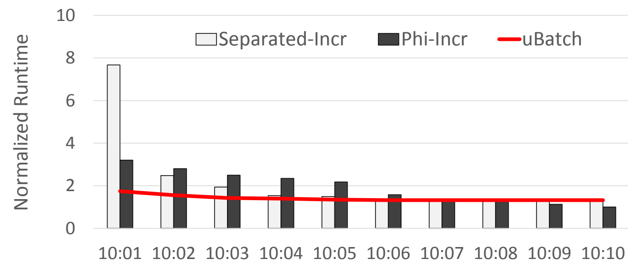


Figure 12. The worst-case perfomance under different capacity ratio between *HE* and *FC* (the capacity of *HE* is fixed)

7.3 Design Space Exploration

We further evaluate our design under different power provisioning capacities. We fix the capacity of heat engine and vary the ratio between heat engine and fuel cell (i.e., adjust the initial installed fuel cell capacity), as shown in Figures 11 and 12. The results shown are the mean value across all traces.

It's clear that over-provisioning improves workload performance. The job turn-around time gradually decreases for all three schemes as we increase the capacity of fuel cell (which increase the total cost of ownership). Due to the cost of additional fuel cell stacks, significant over-provisioning may not be economically attractive. This is especially true when the return-on-investment (ROI) of capacity expansion becomes much lower at high fuel cell capacities. In Figure 11, doubling the installed fuel cell capacity at a *HE/FC* ratio of 10:01 can reduce the job turn-around time by 8% for *Separated-Incr*, while doubling at a *HE/FC* ratio of 10:02 can only yield a further reduction of less than 1%.

Expanding onsite fuel cell generation often have greater influence on the worst-case job performance. For example, by doubling the fuel cell capacity at a *HE/FC* ratio of 10:01 (which increases the total onsite engine cost by 60%) can reduce job turn-around time by 68% for *Separated-Incr*. This number becomes much lower for *Phi-Incr* (12.5%) since *Phi-Incr* has better overall energy and power utilization even if the installed fuel cell capacity is low. The average job turn-around time increase is less than 1% for all the three evaluated designs and can actually be ignored.

Furthermore, by comparing the performance of different power provisioning schemes, we can conclude the following results. (1) μBatch could enable a fuel powered data center to perform better in an energy/power constrained environment. The job turn-around time of μBatch is normally better than both *Separated-Incr* and *Phi-Incr* when the installed fuel cell capacity is low (i.e., less than 30% of the heat engine capacity in Figures 11 and 12). This is because the less responsive power system installed onsite, the higher requirement on the system resiliency of the data center. (2) When the fuel cell is over-provisioned, the 3-step heuristic optimization of μBatch could be inefficient. In Figure 11, μBatch yields less than 1% performance difference compared to *Phi-Incr* at a HE/FC ratio of 10:10, whereas the difference can reach 33% in Figure 12. This is because our intuitive idea of μBatch design is to save the fuel cell capacity for handling power emergency, which cause less-effective utilization of fuel cell when it is over-provisioned. As new materials and power technology emerges, fuel cells could be more cost-competitive in the future. At that time, a straightforward incremental job dispatching scheme such as *Phi-Incr* may be sufficient and better. However, over-provisioning fuel cells is not likely today since they are still one of the most expensive generators.

8. CONCLUSIONS

We expect clean fuel powered data center to be a very promising design practice today and tomorrow, since the energy crisis and climate change have become a growing concern. However, several challenges need to be addressed to realize this vision. First, the cost of conventional power provisioning infrastructure is very high. Second, a coordination scheme between the computing system and smart energy system is lacking.

In this study we propose to use a hybrid and islanded power provisioning architecture, referred to as Phi, to fundamentally cut the cost of a fuel powered data center. We show that the key challenge to transition to such an unconventional design is a "power lagging" issue incurred in today's onsite generators. To solve this problem, we also propose a novel supply-/load -aware load dispatching schemes called μBatch. It allows data center to develop the resiliency to handle power lagging and greatly facilitate the deployment of fuel-based generation onsite. Using realistic data center workload traces, we show that our proposed scheme can reduce data center power provisioning cost (OpEx + CapEx) by over 90% compared to conventional utility power based design, and by over 80% compared to a non-optimized fuel-powered data center. The proposed design seeks a synergism of power generation scheduling and computing load dispatching, and therefore does not significantly affect workload performance. Our results show that the proposed design maintains a near-optimal (within 1%) performance in terms of average job turn-around time, and a significantly lowered worst-case design penalty compared to other design alternatives. The proposed design allows fuel powered data center to aggressively and efficiently utilize its onsite power system in a power/energy constrained environment.

ACKNOWLEDGEMENT

We thank the anonymous reviewers for their valuable comments. Chao Li is supported in part by Shanghai Jiao Tong University New Faculty Start-up Funding and SJTU-MSRA Faculty Award. This work is also partly supported by the National Natural Science Foundation of China (Grant No. 61202026 and No. 61332001) and Program of China National 1000 Young Talent Plan.

REFERENCES

1. The Smart Grid: An Introduction, *White Paper*, U.S. Department of Energy
2. S. Chowdhury, S.P. Chowdhury, and P. Crossley. Microgrids and Active Distribution Networks, *Institution of Engineering and Technology*, 2009
3. The Role of Distributed Generation and Combined Heat and Power (CHP) Systems in Data Centers. *Technical Report*, U.S. EPA. 2007
4. http://maps.nrel.gov/biomass

5. Biomass Holds the Promise of Clean Power and Fuel - If Handled Correctly. *White Paper*. 2012. http://www.ucsusa.org/
6. J. Koomey. Growth in Data Center Electricity Use 2005 to 2010. Analytics Press, 2009.
7. Wyoming Waste Powers Microsoft Datacenter, 2014 http://www.enterprisetech.com/2014/11/06/wyoming-waste-powers-microsoft-datacenter/
8. Energy Server Architecture. Bloom Energy Corporate http://www.bloomenergy.com/fuel-cell/energy-server-architecture/
9. Apple Data Center Does Fuel Cell Industry a Huge Favor, 2012 http://www.technologyreview.com/view/508421/apple-data-center-does-fuel-cell-industry-a-huge-favor/
10. EBAY Hosts Unveiling of Groundbreaking New "Bloom Box" Fuel Cell, 2010 http://blog.ebay.com/ebay-hosts-unveiling-of-groundbreaking-new-bloom-box-fuel-cell/
11. R. Sharma, T. Christian, M. Arlitt, C. Bash, and C. Patel. Design of Farm Waste-Driven Supply Side Infrastructure for Data Centers, *ASME 2010 the 4th Int. Conf. on Energy Sustainability*, 2010
12. P. Banerjee, C. Patel, C. Bash, A. Shah, and M. Arlitt. Towards a Net-Zero Data Center. *Journal on Emerging Technologies in Computing Systems*, 2012
13. An introduction to spark spread, U.S. Energy Information Administration, 2013, http://www.eia.gov/todayinenergy/detail.cfm?id=9911
14. Natural Gas Pipeline System: Reliability and Resiliency. *White Paper*. Bloom Energy Corporate, 2013
15. D. Chiesa and S. Zirkelbach. Microgrids Help More than They Hurt. *IEEE Electrification Magazine*, 2014
16. Schenider Electric Expands in Secured Power. *White Paper*, Schneider Electric, 2003
17. A. Riekstin, S. James, A. Kansal, J. Liu, E. Peterson. No More Electrical Infrastructure: Towards Fuel Cell Powered Data Centers. *Workshop on Power-Aware Computing* (*HotPower)*, 2013
18. C. Vyas and B. Gohn. Smart Grid Consumer Survey, Pike Research, 2012
19. G. Saur, J. Kurtz, C. Ainscough, and M. Peters. Stationary Fuel Cell Evaluation. Technical Report, *National Renewable Energy Lab*. 2014
20. Cost-Effectiveness of Distributed Generation Technologies. *White Paper*, *CPUC Self-Generation Incentive Program*. 2011
21. Onsite Distributed Generation Systems for Laboratories. *White Paper*, U.S. Environmental Protection Agency and U.S. Department of Energy. 2011.
22. Gas Turbine Prices per KW. http://www.gas-turbines.com/trader/kwprice.htm
23. Improving the Operating Efficiency of Microturbine-Based Distributed Generation at an Affordable Price. *White Paper*, U.S. Department of Energy. 2014 http://www.energy.gov/sites/prod/files/2014/09/f18/capstone.pdf
24. Doing the Match on Bloom Energy http://www.forbes.com/2010/02/25/fuel-cell-costs-technology-ecotech-bloom-energy.html
25. W. Turner and K. Brill, Cost Model: Dollars per kW plus Dollars per Square Foot of Computer Floor. *White Paper*. Uptime Institute, 2008
26. Understanding Electric Demand. *White Paper*, National Grid, 2012 https://www.nationalgridus.com/niagaramohawk/non_html/eff_elec-demand.pdf
27. Residential Demand Rate, http://www.georgiapower.com/pricing/residential/residential-demand.cshtml
28. Electricity Energy Storage Technology Options. *White Paper*, Electric Power Research Institute, 2010
29. L. Zhao, J. Brouwer, S. James, J. Siegler, E. Peterson, A. Kansal, and J. Liu, Servers Powered by a 10KW In-Rack Proton Exchange Membrane Fuel Cell System, *ASME 2014 the 8th Int. Conf. on Energy Sustainability*, 2014
30. Z., Zhi, F. Liu, B. Li, B. Li, H. Jin, R. Zou, and Z. Liu. Fuel Cell Generation in Geo-Distributed Cloud Services: A Quantitative Study. *Int. Conf. on Distributed Computing Systems (ICDCS)*, 2014
31. C. Li, R. Zhou, and T. Li. Enabling Distributed Generation Powered Sustainable High-Performance Data Center. *Int. Symp. on High-Performance Computer Architecture (HPCA)*, 2013
32. N. Sharma, S. Barker, D. Irwin, and P. Shenoy. Blink: Managing Server Clusters on Intermittent Power. *Int. Conf. on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, 2011
33. I. Goiri, R. Beauchea, K. Le, T.D. Nguyen, M. Haque, J. Guitart, J.Torres, and R. Bianchini. GreenSlot: Scheduling Energy Consumption in Green Datacenters. *Int. Conf. for High Performance Computing, Networking, Storage and Analysis (SC)*, 2011
34. C. Li, A. Qouneh, and T. Li. iSwitch: Coordinating and Optimizing Renewable Energy Powered Server Clusters. *Int. Symp. on Computer Architecture (ISCA)*, 2012
35. N. Deng, C. Stewart, D. Gmach, M. Arlitt, and J. Kelley. Adaptive Green Hosting. *Int. Conf. on Autonomic Computing (ICAC)*, 2012
36. P. Gao, A. Curtis, B. Wong, and S. Keshav. It's Not Easy Being Green. *ACM SIGCOMM*, 2012
37. C. Li, Y. Hu, R. Zhou, M. Liu, L. Liu, J. Yuan, T. Li. Enabling Datacenter Servers to Scale Out Economically and Sustainably. *Int. Symp. on Microarchitecture (MICRO)*, 2013
38. I. Goiri, W. Katsak, K. Le, T. Nguyen, and R. Bianchini. Parasol and GreenSwitch: Managing Datacenters Powered by Renewable Energy. *Int. Conf. on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, 2013
39. C. Li, R. Wang, T. Li, D. Qian, and J. Yuan. Managing Green Datacenters Powered by Hybrid Renewable Energy Systems. *Int. Conf. on Autonomic Computing (ICAC)*, 2014
40. L. Liu, C. Li, H. Sun, Y. Hu, J. Xin, N. Zheng, and T. Li. Leveraging Heterogeneous Power for Improving Datacenter Efficiency and Resiliency. *IEEE Computer Architecture Letters*, 2014
41. K. Rajashekara. Hybrid Fuel-Cell Strategies for Clean Power Generation, *IEEE Trans. on Industry Applications*, Vol. 41, Issue 3, 2005
42. A. Choudhury, H. Chandra, and A. Arora. Application of Solid Oxide Fuel Cell Technology for Power Generation—A Review. *Renewable and Sustainable Energy Reviews*, Volume 20, 2014
43. S. Govindan, D. Wang, A. Sivasubramaniam, and B. Urgaonkar. Leveraging Stored Energy for Handling Power Emergencies in Aggressively Provisioned Datacenters. *Int. Conf.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, *Int. Symp. on Computer Architecture (ISCA)*, 2011
45. 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
46. Power Generation Technologies. *Chapter 6 of the EPA Combined Heat and Power Partnership*, U.S. Environmental Protection Agency, 2007 http://www.epa.gov/chp/documents/biomass_chp_catalog_part6.pdf
47. http://www.wartsila.com/en/power-plants/learning-center/overview
48. Economics of Fuel Cell Solutions for Backup Power. *White Paper*, Ballard Power Systems Inc., 2012
49. A. Halsey III. Aging Power Grid on Overload as U.S. Demands More Electricity. *The Washington Post*, August 1, 2012.
50. L. Rhodes, US Power Grid Has Issues with Reliability. 2013 http://www.datacenterknowledge.com/archives/2013/01/16/us-power-grid-has-issues-with-reliability/
51. Yank R. Singh, D. Irwin, P. Shenoy, K. Ramakrishnan. Yank: Ena-bling Green Data Centers to Pull the Plug. *USENIX Symp. on Networked System Design and Implementation (NSDI)*, 2013
52. T. Lipman. Overview of Stationary Fuel Cell Systems. *Report*, US Department of Energy Clean Energy Application Center, 2010 http://www.ncsl.org/documents/energy/TLipman0910.pdf
53. Data Collection of Diesel Generators in South Australia. *White Paper*, IT Power (Australia) Pty Limited. 2013
54. Logs of Real Parallel Workloads from Production Systems http://www.cs.huji.ac.il/labs/parallel/workload/logs.html
55. B. Lawson, E. Smirni, and D. Puiu. Self-adapting Backfilling Scheduling for Parallel Systems. *International Conf. on Parallel Processing (ICPP)*, 2002.
56. Cost and Performance Data for Power Generation Technologies. *Technical Report*, Black & Veatch Holding Company, 2011