Large-scale and Extreme-scale Computing with Stranded Green Power: Opportunities and Costs

Fan Yang, *Member, IEEE,* and Andrew A. Chien, *Fellow, IEEE*

Abstract—Power consumption and associated carbon emissions are increasingly critical challenges for large-scale computing. Recent research proposes exploiting stranded power – uneconomic renewable power – for green supercomputing in a system called Zero-Carbon Cloud (ZCCloud) [1], [2], [3]. These efforts studied production supercomputing workloads on stranded-power based computing resources, demonstrating their achievable productivity.

We explore economic viability of stranded-power based supercomputing, using three datacenter total-cost-of-ownership (TCO) models to study cost-effectiveness. These studies show that ZCCloud's approach can be cost-effective in the USA today, and is even more attractive in regions with higher power prices (e.g., Japan, Germany), achieving cost advantages as large as 50%. Environmental and power-grid benefits are a further advantage. We also explore the sensitivity of these results to changes in hardware TCO; cheaper hardware or longer lifetimes magnify the attractiveness of stranded-power based approaches, yielding advantages as large as 91%. These results are robust across different TCO models. Finally, we study extreme-scale supercomputers ($>$ 100MW), finding stranded-power can increase peak capability per cost by as much as 80%.

Index Terms—Supercomputing, extreme scale, cost, power grid, stranded power.

1 Introduction

THE growing consensus on climate change due to anthropogenic carbon [4], [5] is driving growing worldwide efforts to reduce carbon emissions [6], [7]. Ambitious "renewable portfolio standards" (RPS) goals for renewable power as a fraction of overall power have been adopted widely across the US (25% by 2025 in Illinois, Minnesota; 50% by 2030 in California, New York [8], [9]), and Obama's "Clean Power Plan," (August 2015) targets 32% reduction in electric power carbon emissions by 2030.

The rapid growth of intermittent renewable generation around the world is producing growing quantities of curtailment (discarded) and uneconomic (negative price) power that we collectively term *stranded power*. Stranded power has been documented as a large, growing untapped resource. For example, in 2014 the Midcontinent Independent System Operator (MISO) power grid curtailed 2.2 terawatt-hours (TWh) of power (see Figure 1) and bought 5.5 TWh at negative price for a total of 7.7 TWh of stranded power from wind resources [3], [10], [11]. And in China, curtailed wind power has grown to 34 TWh in 2015 [12], suggesting stranded power of 60-100TWh. Around the world, as renewable generation fraction increases due to rising RPS standards, stranded power is projected to increase significantly [11], [13], [14].

Information and computing technologies (ICT) produced carbon emissions (8% of electric power in 2016) are growing most rapidly, and are projected to reach 13% by 2027 [15], [16]. Supercomputers and datacenters are major elements of this consumption. With Dennard scaling long over [17], [18] power levels are growing rapidly: today's 15-petaflop systems are $>$ 10 MW [19], [20], announced near-exascale machines will exceed 15 MW [21], [22], with Exascale systems in 2023 projected $>$ 25 MW. Centers often operate under power "caps", effective carbon footprint limits [23], or forced power reductions [24]. Direct power cost, as well as implied costs in cooling, facilities, and the associated environmental impact [5], [25], [26], is a major challenge for supercomputers at extreme-scale, inspiring radical approaches to scaling datacenters [1], [27], [28].

Here we build on prior research that proposes to power supercomputers *exclusively* with stranded power, called *Zero-Carbon Cloud* (ZCCloud) [1], [2], [3]. Detailed analysis of stranded power and the computing capabilities of ZCCloud are published in [3], [29]. Because stranded power has challenging, irregular availability intervals, these efforts demonstrated that, despite such irregularity, ZCCloud resources can be highly productive for HPC workloads.

A variety of research efforts study addition of local renewable generation to datacenters to reduce brown power usage [30], [31], [32], [33], [34], [35], [36], by exploring workload scheduling, compute management, etc. In operating infrastructure, cloud companies such as Google, Microsoft, and Facebook have made extensive long-term purchase agreements [37] to help the financing of new renewable power generation facilities. ZCCloud differs from these approaches in three ways. (1) ZCCloud exclusively uses stranded power, and does not rely on any stable grid power that are mostly brown, and hence eliminates the incremental carbon footprint from both brown power generation and transmission. (2) By co-locating with the source of stranded power, ZCCloud operators do not need to pay for the renewable generation facilities that already exist, as well as power transmission facilities. (3) ZCCloud resources must respond to and manage the intermittence of power.

- *Fan Yang and Andrew A. Chien are with the Department of Computer Science, the University of Chicago.*
- *Andrew A. Chien is also affiliated with the Mathematics and Computer Science Divsion at Argonne National Laboratory.*

Manuscript received December 2016; revised XX, yy, ZZZZ.

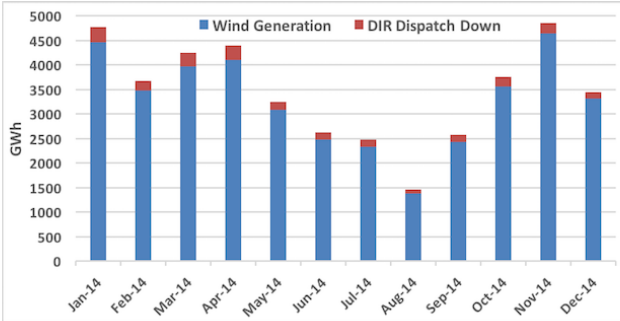


Fig. 1. Wind generation and down dispatch (curtailment) (MISO, 2014). Total of 2.2TWh is a fraction of stranded power in MISO and globally.

Intermittence arises from renewable generation variability, grid operations, and complex energy markets. This gives rise to much greater volatility, including rapid swings in periods as short as 5 minutes, as documented in Section 2.3. Intermittent computing resources have been widely used for decades [38], [39], [40], [41], [42], [43], and successfully applied to large scale applications [44], [45], [46], [47], [48], [49], [50].

While the productivity of ZCCloud is promising, in large-scale computing, cost-effectiveness is critical. So in this paper, first, we apply total-cost-of-ownership (TCO) models, combining them with the productivity results, to assess the cost-effectivenss of stranded-power based computing. Because TCO models can differ by organization, to reduce sensitivity to assumptions, we employ leading academic and industrial TCO models [51], [52], creating variants for stranded-power based supercomputers that reflect their costs (e.g., reduced power cost, increased external networking). Our goal is to understand that even without considering the environmental and power-grid benefits, whether these lower costs can make up for lower duty factor. Second, because we believe stranded-power based computing resources will be used to complement traditional resources, we study a variety of scaling scenarios, varying system sizes and resource mix to understand tradeoffs. Third, as regional power price circumstances vary today, and future projections vary widely, we compare cost-effectiveness under a variety of power price and compute hardware price scenarios. These studies provide a perspective on the geographies and future scenarios where stranded power based supercomputing is most attractive. Finally we examine extreme scaling at Exascale, and beyond. Specific contributions include:

- Evidence that stranded-power based resources can be cost-effective in today's TCO environment (compute, infrastructure, and USA power price). In short, lower infrastructure and power costs can make up for lower duty factor; this viability is robust across TCO models with benefit ranging 3-34%. So they can not only reduce carbon emissions, but can be cost-effective.
- Exploring power price across a range of current and future scenarios, the advantage of ZCCloud grows with increasing power price. The three TCO models estimate ZCCloud 16-63% improvement of cost-effectivenss at high power price typical for Germany and Denmark.
- Future hardware prices are uncertain, if significant decreases are seen, ZCCloud's cost-effectiveness advantage improves. For 0.25x hardware price scenario, cost-effectiveness increases 14-91%. In high hardware price scenarios, ZCCloud remains attractive, but its advantage is modest.
- Exploring Exascale and beyond, using Berral's model, suggests ZCCloud can increase cost-effectiveness by 45% and increase peak PFLOPS by up to 80% under a fixed budget.

The remainder of the paper is organized as follows: Section 2 provides background on renewables, power markets and stranded power, extant stranded power, and datacenter cost modelling. Section 3 describbess the ZCCloud approach. We assess productivity and cost-effectiveness in Sections 4, 5, and 6. Next, we consider and extreme scaling study in Section 7, and discuss related work in Section 8. Section 9 summarizes and proposes directions for future work.

2 Background

In this section, we summarize growth of renewable power, energy markets and stranded power, quantitative features of stranded power, datacenter cost modelling.

2.1 Renewable Power Generation

Fossil fuels are major sources of anthropogenic carbon emissions, and are widely believed to be driving climate change [5]. Such concerns drive the rapid development of renewable energy generation. Growth of solar and wind is the most rapid, comprising 5.2% of US power generation in 2014 [53]. In many states renewable power levels of more than 10% have already been achieved. For example, California achieved 20% renewable portfolio standard (RPS) in 2010 [54]. Germany achieved 33% RPS in 2015. More ambitious RPS goals have been adopted: 50% by 2030 in California and New York [8], [9]. President Obama released the "Clean Power Plan" in August 2015 that establishes the national standards to limit carbon pollution from power plants. In 2015, the U.S. Department of Energy Wind Program announced a 35% RPS, 404GW goal for wind by 2050 [55]. In short, RPS's are growing rapidly.

While attractive because they are zero-carbon, wind and solar challenge power grids because of their volatility, on average delivering 30% of their maximum generation (called capacity factor). This volatility creates grid management challenges for reliable power, increasing grid capital costs for generation plants, transmission, and other infrastructure [13], [14], [56]. And in future high-RPS scenarios generation, peak renewable generation will exceed load by significant amounts. For example, in a 50% RPS power grid with 100GW load, a 30% capacity factor would require (50% * 100GW)/30% = 167GW peak renewable generation.

2.2 Power Grid, Energy Markets, and Stranded Power

Modern ISOs dispatch generation and price power generation in real time. For example, the Mid-continent ISO (MISO) market [57] uses 5-minute intervals: generators offer power to the grid, and it prices their offered power with locational marginal price (LMP), which varies by site and depends

TABLE 1
MISO Market Statistics Jan 1, 2013 to Apr 14, 2015 (28.5 months) [3]

MISO	Overall	Wind
Generation Sites	1,259 Total	200 Wind
5-minute Periods	76.9M Total	36.6M Wind
Power	1,188 TWh Total	89 TWh Wind
$$ Revenue	$95B Total	$4B Wind

on transmission structure and supply-demand balance. An important goal for power purchase is "merit order", where lower prices have higher priority. Scheduling power is difficult because generation and demand must be matched instantaneously and in general power is not stored. Any sudden change in demand or supply must be addressed. So, to ensure stability or when there is overgeneration, ramp constraints, or transmission congestion, Power grids will "curtail" (aka, "spillage", or "dispatch down"), preventing power from entering the grid, and power grid markets send localized price signals that power is not useful by pricing the power with negative locational marginal price (LMP).

In response generators will reduce production. If they cannot reduce it or generate for external reasons, they will sell power at a negative price. We define *Stranded Power* (SP) as offered power generation with no economic value, thus including both curtailment and power delivered with zero or negative LMP. Typically, stranded power is a localized phenomenon in space and time, occuring in a few places and areas of the grid. Therefore, stranded power is not in general available, for example, at arbitrary sites chosen for traditional datacenters [29].

Both wind and solar exhibit temporally-correlated generation, creating significant stranded power challenges [13], [58]. Stranded power is a large-scale and growing phenomenon around the world. It has been documented at the scale of terawatt-hours (TWh) across the USA, [10], [54], across Europe (Denmark, Germany, Ireland, and Italy) [56], and in China stranded wind power exceeded 34 terawatt-hours (TWh) in 2015 [12]. As shown in Figure 1, the monthly wind generation and dispatch down of Mid-continent Independent System Operator (MISO) for 2014 shows that the total "dispatch down" was 2.2 TWh (183 MW average), 7% of wind generation. Analysis including negative-priced power yields 7.7 TWh for stranded power in 2014. A detailed, 3-year study of MISO confirms stranded power as a persistent and growing phenomenon [29]. With aggressive increases in renewable portfolio standard (RPS) goals from 10 to 33-50%, many projections suggest rapid grown of stranded power [8], [13], [14].

2.3 Stranded Power Quantified

A number of studies of curtailed power (a subset of stranded power) characterize quantity at long time scales (months or years) [10], [12], [56]. A recent detailed study [29] characterizes stranded power at 5-minute temporal resolution, and characterizes its temporal and geographic distribution by analyzing MISO market data [57] (summarized in Table 1).

As defined in Section 2.2, Stranded Power is power generation that with no economic value. We use two definitions LMP (instantaneous) or NetPrice (average over a longer interval), and an average price threshold of $0 up to $5 determine whether the power is stranded or not. These two models are formally defined below, and are discussed extensively in our previous studies [3], [29] that characterize the quantity (MW or MWh) and temporal distribution (*duty factor*, the fraction of time when stranded power is available). Based on highest duty factor, we select the best generation sites for stranded power.

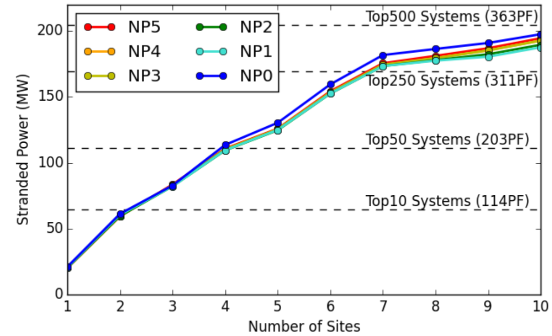


Fig. 2. Average Stranded Power generation at highest duty-factor wind Generation sites (NetPrice).

1) LMPc: LMP$<c$
($c = \$0, \$1, ..., or \$5$) for a model.

2) NPd or NetPrice$d$: NetPrice$<d$
($d = \$0, \$1, ..., or \$5$) for a model.

$$NetPrice = \frac{\sum_{period} LMP \cdot Power}{\sum_{period} Power} \quad (1)$$

Quantity: Our previous studies show that within MISO, a major US ISO, the quantity of stranded power is sufficient to power the world's largest 250 supercomputing systems [59] (see Figure 2). For all of the NetPrice (NP) scenarios, two sites are enough for the Top 10 systems, and seven sites are enough for the Top 250 systems. There are more than 5 grids of this scale in the USA, and more than 30 worldwide. Furthermore more as renewable generation grows, stranded power is expected to increase. So, we expect that the quantity of stranded power will not be a practical limit for ZCCloud in the next several decades.

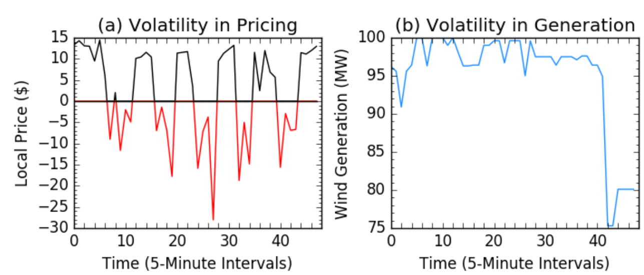


Fig. 3. Local Price can swing from -28/MWh to $+9$/MWh (dollars) in 5-minutes, despite steady renewable generation (MISO Generator 6414, 2/28/2016)

Temporal Distribution (intermittence): Studies show that the temporal distribution of stranded power, because it includes market and grid operations effects, is much more volatile than renewable generation alone. Figure 3 shows that price swings across positive and negative values – often as fast as in 5-minutes – despite steady renewable

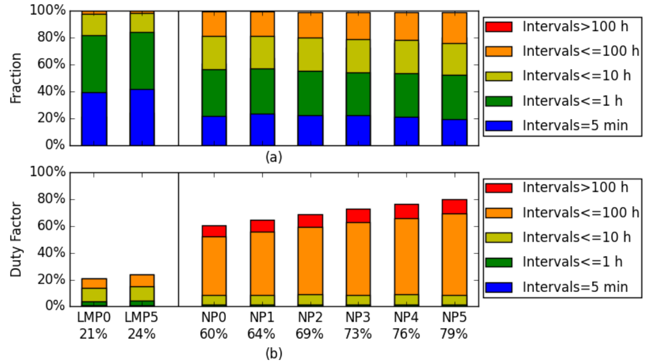


Fig. 4. (a) Stranded Power Interval size distribution, (b) SP interval contribution to duty factor, by size. Percentages indicate the achieved duty factor.

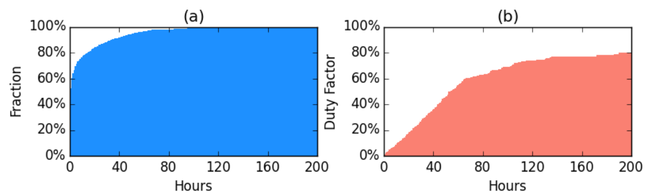


Fig. 5. (a) CDF of Stranded Power Interval size (NP5), (b) CDF of SP interval contribution to duty factor, by size (NP5).

production. This changes faster and is less predictable than most large-scale solar and wind generation fluctuations.

Using the stranded power models, the temporal distribution varies by model, but NetPrice produces many long intervals. Because stranded power only exists in large quantities on a small number of locations in the grid, we study some of the best sites. Digging deeper, analyzing one of the best stranded power sites, Figure 4(a) shows the distribution of SP interval durations. For LMP, 70% are shorter than one hour, and overall duty factors are less than 30%. For NetPrice half of the SP interval durations are >1 hour, and those intervals account for a 50-70% contribution to duty factor (see Figure 4(b)). NetPrice effectively masks the market price fluctuations, and achieves 60% to 79% duty factor, but cannot eliminate long periods without stranded power (over 500 hours). Figure 5 shows the detailed CDF for NP5. Most of SP intervals are shorter than 80 hours, making up an >60% duty factor. There are a few long intervals (100's of hours) that suit long-running jobs.

In summary, stranded power is available in great quantity and sufficient to power today's supercomputers of 10's of MW at a single site, with power available in long intervals for up to 79% of the time. And, numerous such sites exist.

2.4 Datacenter Cost and Modelling

Today's peta-scale supercomputers are comprised of thousands of servers, requiring 10's of megawatts of electricity, and thousands of sqft. To accommodate, datacenters provide physical space, cooling, and power. Building and operating datacenters costs millions of dollars per year [21], [22], [60]. As a result, managing TCO is a major operational and research effort. Dominant elements of cost are compute servers and datacenter facilities, together comprising more than 75% of total cost of ownership (TCO) [61], [62]. Other cost factors include power, system administration,

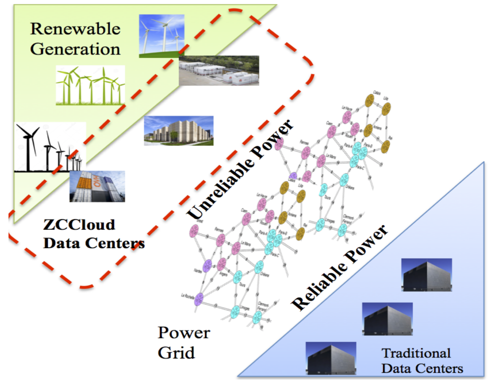


Fig. 6. ZCCloud datacenters exploit unreliable stranded power at generation, providing no load to the power grid.

power transmission, and data networking. Widely used TCO models for research include those published by leaders at Google [61], Amazon [52] and academics now associated with Microsoft [51], [63], [64].

3 Zero-Carbon Cloud (ZCCloud): Supercomputing with Stranded Power

We study Zero-Carbon Cloud (ZCCloud), a new approach that exploits stranded power, power that is uneconomic in the power grid. As shown in Figure 6, ZCCloud datacenters use only unreliable stranded power, so they are logically on the generation side of the grid, and would typically be attached at renewable power generation site or grid transmission node.

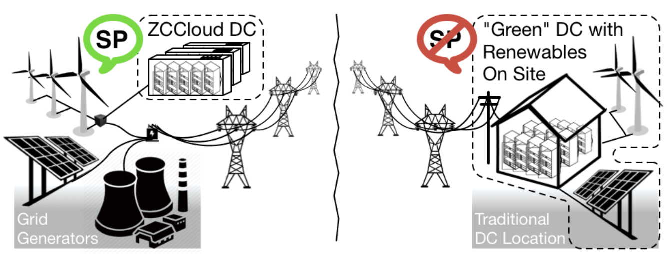


Fig. 7. ZCCloud datacenters are placed at locations with stranded power (typically at renewable generators). Traditional datacenters typically have no access to stranded power, and building on-site renewable generation requires paying for such facilities (circled by dashed line).

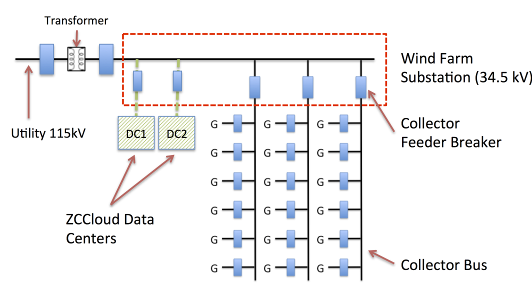


Fig. 8. ZCCloud datacenters can attach to wind farm substations, providing access to farm-wide power (100's MW) and isolation from both grid and individual generators (G). 34.5 kV connection to Datacenter main switch boards then internal HVDC distribution [65].

Advantages for the ZCCloud approach: (1) ZCCloud absorbs low (or zero) cost stranded power that would otherwise have been wasted. Therefore, it does not compete with normal grid loads, and can increase the grid's ability to accept renewable power [14]. It also reduces power costs for the datacenter. (2) ZCCloud only exploits uneconomic power, requiring no incremental increase in generation and thus no increase in carbon footprint. Elimination (not offset) of the power consumption carbon footprint of the datacenter is a significant step beyond current operations and future plans of any large-scale cloud provider [66], [67]. (3) ZCCloud exploits renewable generation that already has an economic relationship with the grid that pays for them, and thus ZCCloud does not incur costs for wind turbines or solar panels (see Figure 7). (4) ZCCloud is typically co-located with existing renewable generation facilities, reducing power transmission and distribution costs, and leveraging the distributed power buses within a wind farm. It can also eliminate building costs, and enable aggressive free-cooling. Figure 8 depicts on possible approach (see DC1, DC1) for wind farm integration. 34.5 kV integration eliminates additional transformers and additional breakers typically required.

Disadvantages of the ZCCloud approach include that resources are only active when stranded power is available, reducing productivity and requiring the computing workload to tolerate intermittence.

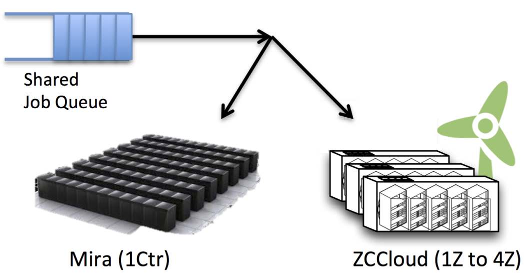


Fig. 9. System overview of Mira-ZCCloud.

Since ZCCloud resources are intermittent, we deploy them as a complementary extension of traditional resources. In subsequent sections, we model configurations that logically combine ZCCloud resources with a traditional supercomputing system such as Mira (see Figure 9). We couple them together using a high-speed network to share a single filesystem, workload and scheduler. Jobs submitted to the shared queue are seamlessly scheduled to the ensemble. During ZCCloud uptime, submitted jobs are assigned to either traditional resources or ZCCloud resources. When ZCCloud is down, jobs go only to the traditional supercomputing resources.

Thus the key questions for ZCCloud viability are:

- How does intermittence affect resource productivity? what are the key parameters of capabilities?
- How does exploiting stranded power affect the total cost of ownership (TCO) of a datacenter?
- When can TCO benefits of ZCCloud overcome the disadvantate in duty factor to achieve competitive cost-efficiency?
- At extreme performance and power levels, what opportunities and new capabilities does stranded power offer?

We address these questions in the following sections.

4 Productivity of Intermittent Resources

To assess the performance benefit of adding intermittent computing resources to traditional datacenter, we simulate HPC workloads first with a daily periodic model, and then with a more complex set of uptime intervals derived from a single wind generation site's stranded power – derived from real MISO power grid statistics. We vary system configurations, and compare system throughput.

4.1 Methodology

System Model We use ALCF's Mira system [60] as a model for traditional datacenter resources, a 10 PFlop IBM BG/Q system with 49,152 nodes, 786,432 cores, and 768 TB memory. Intermittent resources (ZCCloud) are connected to datacenter (*Ctr*) as an extension, denoted *Ctr+Z*. We compare Ctr+Z with datacenter-only system (*Ctr*). Moreover, we consider ZCCloud extension at varied scale: 1x, 2x, and 4x of Mira resources, denoted *Ctr+1Z, Ctr+2Z, and Ctr+4Z*. For fair comparison, we also scale datacenter resources to 2x, 3x, and 5x of Mira, denoted *2Ctr, 3Ctr, and 5Ctr*, and use *1Ctr* as baseline. Note that Ctr+{n}Z is of the same size as {n+1}Ctr.

Job Scheduling and Workload We simulate ALCF job trace [60] using Mira's job scheduler, Cobalt v0.99 integrated with a simulator, Qsim [68]. Workload properties are summarized in Table 2. Workloads are scaled to match utilization on Mira, adding jobs with the same distributions of attributes (job size is not increased). The scheduler assigns jobs equally to datacenter and ZCCloud intermittent resources when ZCCloud is available, and only to the datacenter when ZCCloud is shutdown. Since the NetPrice models produce stranded-power dominated by intervals of 10 hours or more (longer than most runs), the scheduler can be highly productive. Because we believe intervals can be predicted (current grids use 7-day and day-ahead forecasts of growing accuracy, and day-ahead markets), we inform the job scheduler of the duration of each ZCCloud interval, so it only schedules jobs that complete before shutdown. Our TCO models include the cost of hardware to capture job state if the interval is mispredicted.

Intermittent Resource Models Two resource models: (1) Periodic intermittent resources: power is available in a fixed

TABLE 2
ALCF Workload Trace

Parameters	Values
# Jobs	78,795
Time Period	12/31/2013—12/30/2014
Runtime (hrs)	0.004—82, Avg 1.7, StDev 3.0
# Nodes	1—49,152, Avg 1,975, StDev 4,100
Rsc Utilization	84% of Mira (100% availability)
App Domains	Physics, Chemistry, Material Science, CompScience, Biology, Engineering

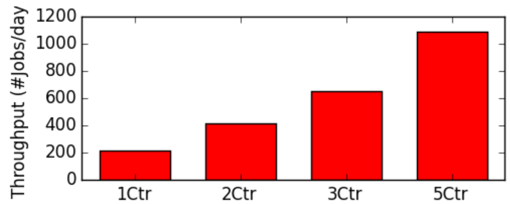


Fig. 10. Throughput vs. System size for Ctr.

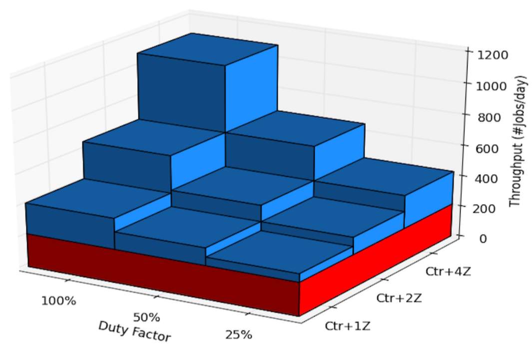


Fig. 11. Periodic Resources: Throughput vs. Duty factor vs. System size for Ctr+Z (Blue) and Throughput of 1Ctr (Red).

daily cycle, (e.g., 8:00 to 20:00 each day), and (2) Irregular intermittent resources: single-site stranded power history from MISO [57] as described in Section 2.3. To normalize between them, we use *duty factor*, the fraction of time when resources are available. For example, a periodic model with uptime 8:00-20:00 has 50% duty factor, and an irregular model using LMP0 has 21% duty factor.

Metrics We use *throughput* (jobs per day) to quantify system productivity. Our prior work [3] also captured the delay in job completion time caused by intermittence.

4.2 Productivity of Periodic Resources

We first explore the scaling of traditional datacenter resources (see Figure 10). The system throughput of Ctr scales linearly with resources, reaching throughput of 215 jobs/day (1Ctr), and 1085 jobs/day with 5x resources (5Ctr).

Next, we evaluate the benefits of periodic intermittent resources, varying duty factor and system size (Figure 11). For periodic resources, both increasing duty factor and adding intermittent resources improve throughput in the same proportion. For example, {Ctr+1Z, 50% duty factor} and {Ctr+2Z, 25% duty factor} achieve similar throughput, both providing 1.5x of 1Ctr node hours. At 100% duty factor, the ZCCloud resources are always available, so the capability of Ctr+{n}Z matches {n+1}Ctr.

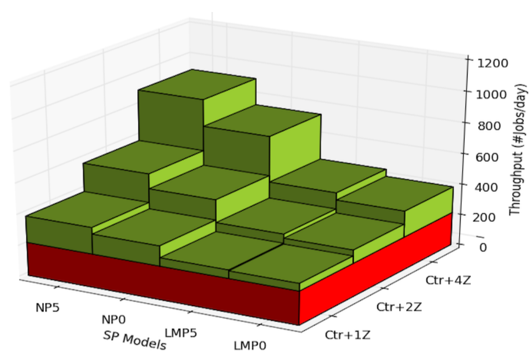


Fig. 12. Irregular Resources: Throughput vs. SP model vs. System size for Ctr+Z (Green) and Throughput of 1Ctr. (Red)

4.3 Productivity of Irregular Resources

We consider four cases of irregular resources, using different SP models including LMP0, LMP5, NetPrice0, and NetPrice5. As we have described in Section 2.3, the corresponding duty factors are 21%, 24%, 60%, and 79% respectively.

Figure 12 shows the throughput of irregular resources for different SP models and system scales. Because it has much more computing hardware, Ctr+Z achieves better throughput than 1Ctr for all cases, but as the SP model provides higher duty factor (such as NetPrice >60%), the advantage grows. For example, at the scale of Ctr+1Z, NetPrice5 provides 1.8x of 1Ctr node hours while LMP5 only provides 1.24x. Throughput scales linearly with resources size and duty factor because most jobs are smaller than the system scale and interval duration. For example, while job runtime is 1.7 hours in average and no longer than 82 hours, NetPrice5-based resources can provide uptime much longer than 100 hours. Overall the irregularity of intervals has little impact, throughput is very similar to comparable periodic resources at the same duty factor.

Overall, adding intermittent resources to datacenter improves throughput, and the benefit is determined by duty factor and scale. Irregular intermittent resources at high duty factors (>50%) give comparable benefit as periodic resources.

5 Modeling Resource Cost

Given supercomputing capability can be scaled up using intermittent resources, here we consider overall costs of ZCCloud systems. We quantify the costs of both datacenter and ZCCloud resources with three total-cost-of-ownership (TCO) models, applying them for a broad range of power price and hardware price scenarios.

5.1 TCO: Traditional and ZCCloud

The Total Cost of Ownership (TCO) of datacenters includes several elements: (1) computing hardware (e.g., servers, racks, internal networking), (2) datacenter physical facilities (DCF), e.g., building, raised floor, power distribution, maintenance, etc., (3) power,and (4) external networking, the cost of Internet connection. Note that we model external internet cost including only fiber installation, and excluding right-of-way costs [1]. Thus,

$$TCO(n) = n \cdot (C_{compute} + C_{DCF} + C_{power})) + C_{net} \quad (2)$$

These costs are parameterized as annual operating cost per system unit, using n to denote the number of units.

1. Right-of-way costs are difficult to estimate, but we conjecture that such costs in the remote areas with large amounts of stranded power are low on a per mile basis. The alignment of ZCCloud data centers with power grid infrastructure may also produce cost benefits in physical infrastructure and right-of-way for network connection.

TABLE 3
Parameter Values in TCO Models

Parameters	Berral's Model	Hamilton's Model
Unit of System	4MW Mira (IBM BG/Q) [60]	9.28MW avg load Datacenter
$C_{compute}$	\$21M/unit, quotes from Mira [60]	\$27.5M/unit for 3 yr amort, \$18M/unit for 5 yr amort
C_{DCF}	\$21M/unit (C_{DCF}=$C_{compute}$ based on [61])	\$9.2M/unit
C_{power}	\$2.1M/unit (US power price, \$60/MWh [69])	\$5.7M/unit (power price=\$70/MWh)
$C_{z,DCF}$	\$2M/unit (Vendor quote in [1])	\$1.6M/Unit
C_{net}	\$0.8M/unit from [70], [71]	
C_{SSD}	\$0.3M/unit (Intel SSD [72])	\$0.7M/unit (same)
$C_{battery}$	\$0.1M/unit (Tesla PowerWall [73])	\$0.2M/unit (same)
C_{cool}	\$0.3M/unit (Free cooling HW [51])	\$0.8M/unit (same)

ZCCloud's TCO differs from traditional datacenters in three ways: (5) additional SSD and energy storage to checkpoint-restart jobs interrupted by a power outage.[2] We include cost for minimal energy storage (e.g., UPS) and SSD's to capture job state. (6) reduced DCF cost, by using containers and colocating at generation sites and avoiding power distribution costs. This reduces physical facility costs significantly. (7) additional hardware for free cooling, and (8) ZCCloud uses stranded power, which is power that has no economic value in the power grid. We assume that this power can be used *at zero cost*, as this is often a better deal than the negative prices that wind generators suffer. Thus, the adjusted ZCCloud TCO is follows:

$$TCO_z(n) = n \cdot (C_{z,compute} + C_{z,DCF} + C_{cool}) + C_{net} \quad (3)$$

$$C_{z,compute} = C_{compute} + C_{SSD} + C_{battery} \quad (4)$$

where $C_{z,DCF}$ is the cost of container per unit, and C_{cool} denotes the free cooling hardware cost, amortized.

TCO models differ by organization, so we consider leading academic and industrial TCO models variants [51], [52].
Berral's Model (Berral) We define a simple TCO model, adapting Berral's model [51], simplifying it by omitting elements that contribute less than 1% to TCO (see Appendix A). We then use a Mira BG/Q system as the unit of scaling with attributes of 4MW, 10PFlops, and cost of \$100M. For example, if $n = 2$, such as 2Ctr, the system is 8MW. Facility cost is simplified based on discussions with the Argonne LCF team [74] and published estimates [61]. Parameters are shown in Table 3. The traditional datacenter and ZCCloud costs are then estimated as above (see Section 5.1).
Hamilton's Model (Hamilton3 and Hamilton5) Hamilton's TCO model [52] is widely used in the industry and cited in academic work. It has the same basic elements: computing and networking hardware, facilities (including cooling), and power. Hamilton's model comprises compute hardware with lower power-efficiency, lower facilities cost per megawatt, and short compute hardware lifetimes. Consequently, compute hardware and power account for a greater fraction of TCO. Values for Hamilton's TCO model are shown in Table 3.

To compare cost models fairly with Berral, we should use models from systems of comparable scale and capability. Hamilton's model system has 50K nodes; our Berral model for Mira has 49K. We estimate Hamilton's base system at approximately 5 petaflops peak, one-half of Mira's 10 petaflops (both in 2011). We deem these systems roughly equal scale, and compare the cost models (NOT the systems) in the remaining sections of the paper.

When applying Hamilton's model to ZCCloud, we eliminate the cost of power distribution and cooling from datacenter facility costs to get the cost of ZCCloud facilities. Hamilton's model assumes compute hardware is amortized over three years, denoted as *Hamilton3*. Most HPC facilities operate machines for at least five years, so we define a variant (*Hamilton5*) that uses a 5 year server amortization.

5.2 TCO for ZCCloud Systems

In Figure 13, we illustrate the three TCO models for the baseline costs of one system unit for both traditional datacenter resources and ZCCloud intermittent resources. In all cases, the largest component of TCO is for compute hardware, and the reduced physical infrastructure and power costs are the reason for ZCCloud resources TCO advantages of 45%, 27% and 34% in Berral, Hamilton3 and Hamilton5 respectively. Berral's model gives a larger advantage to ZCCloud due to the larger cost of facilities in Berral than in Hamilton. Because external networking, SSD, energy storage, and free cooling costs are relatively small, we group them under "Other" in this graph and all subsequent ones.

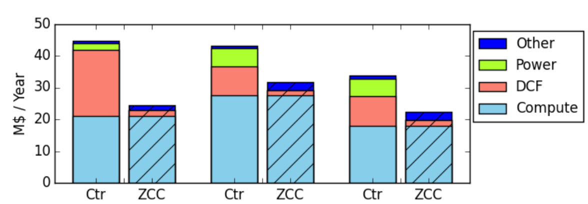


Fig. 13. TCO of traditional resources and ZCCloud intermittent resources.

2. This minimal energy storage (and SSD's) enables checkpointing jobs when stranded power ends; we use a standard checkpoint and restart.

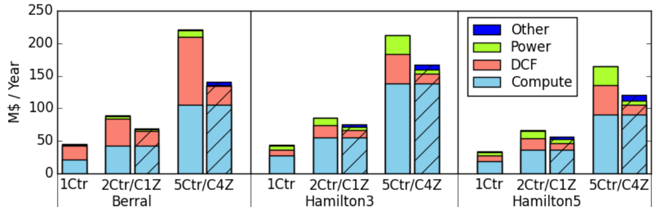


Fig. 14. TCO of Ctr and ZCCloud at increasing scale. We compare {n+1}Ctr and Ctr+{n}Z (denoted as C{n}Z).

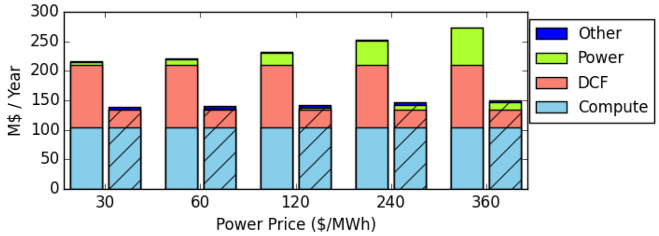


Fig. 15. Berral's TCO vs. Power price. (Left bars for 5Ctr, right bars for Ctr+4Z.)

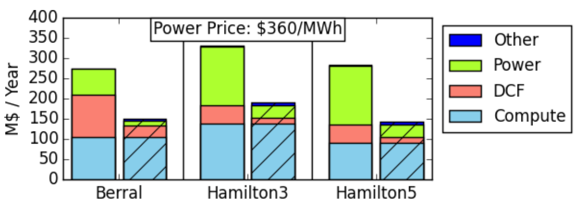


Fig. 16. Three TCO models for power price=$360/MWh. (Left bars for 5Ctr, right bars for Ctr+4Z.)

Considering scaling produces the TCO's for traditional datacenters and ZCCloud shown in Figure 14. The TCO reduction of ZCCloud systems grows with increasing system scale, as the higher TCO for the one reliable Ctr unit is amortized over more resources. This trend is seen in all three TCO models, though the net advantage is smaller in Hamilton's models. At the right, Ctr+4Z (C4Z) is 37%, 21%, and 27% cheaper than 5Ctr in Berral, Hamilton3 and Hamilton5 respectively.

5.3 TCO vs Power Price

Power price varies widely by country, region, and over time. [75]. For example, the average retail price for power in Germany in 2011 was $350/MWh, and in Denmark, the greenest country in Europe, $410/MWh. On the other side of the globe, Japanese power costs $250/MWh. We use $60/MWh as the US power price, which is close to $70/MWh in Hamilton's model. For completeness, we consider a range down to the wholesale prices in the US, which are $30/MWh for the Midcontinent ISO (MISO).

In Figure 15, we study TCO for Berral's model varying power price $30-$360/MWh for 5Ctr and Ctr+4Z systems. At higher prices, power cost increases as a fraction of TCO. At $360/MWh, the Ctr+4Z system achieves 45% lower TCO, and at the lowest power price ($30/MWh), the advantage is reduces to 35%.

Looking across models in Figure 16 shows an interesting story, Hamilton's models project a large TCO reduction at the highest power price, because it estimates power costs as a larger fraction, it achieves even larger cost reductions

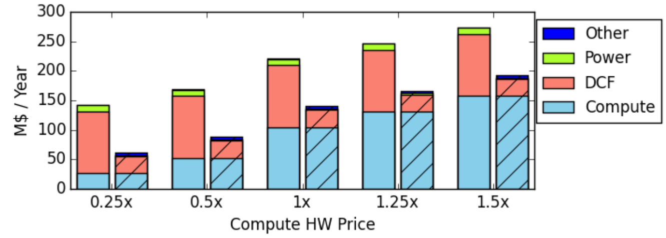


Fig. 17. Berral's TCO vs. Compute HW price. (Left bars for 5Ctr, right bars for Ctr+4Z.)

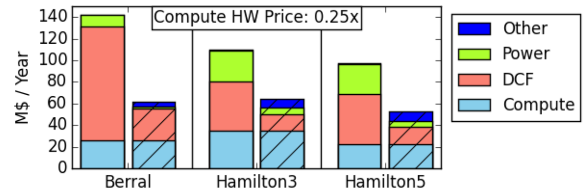


Fig. 18. Three TCO models for compute HW price=0.25x. (Left bars for 5Ctr, right bars for Ctr+4Z.)

than in Berral, 42% and 49% for Hamilton3 and Hamilton5 respectively. We will see the impact of this in overall cost-effectiveness in Section 6.

5.4 TCO vs Compute HW Price

Next we consider how changing compute hardware prices affect ZCCloud benefits. At the End of Moore's Law [18], [76], hardware price may change significantly – due to longer server lifetimes or commoditization (Intel processors at ARM prices). On the other hand, if slow progress continues, with heroic multi-patterning lithography, prices could creep upward, perhaps 1.5x [77]. So, we consider a range of prices.

We explore Berral's TCO model, varying the compute hardware price (see Figure 17) from 0.25x to 1.5x. Hardware cost is the largest component of datacenter TCO. Lower compute hardware price increases the relative advantage of ZCCloud, because hardware cost is more important than in traditional systems. With computing hardware price 0.25x of current level, Ctr+4Z TCO is 57% cheaper than 5Ctr. At high computing prices, ZCCloud benefits are reduced, only 30% lower than traditional.

Figure 18 compares three TCO models for the 0.25x hardware price scenario. Hamilton's models are hardware cost heavy, so reducing hardware price make them more similar to Berral's and magnifies the importance of the cost that ZCCloud can reduce (facilities, power, etc.). The results show ZCCloud TCO benefits of 41% and 46% in Hamilton3 and Hamilton5 respectively.

All three TCO models show that ZCCloud can reduces total cost for computing capability significantly, with a growing advantage if power becomes more expensive or hardware TCO decreases.

6 Exploring Cost-Performance for ZCCloud

While employing ZCCloud resources reduces costs per unit, the frequent downtime of stranded-power based resources

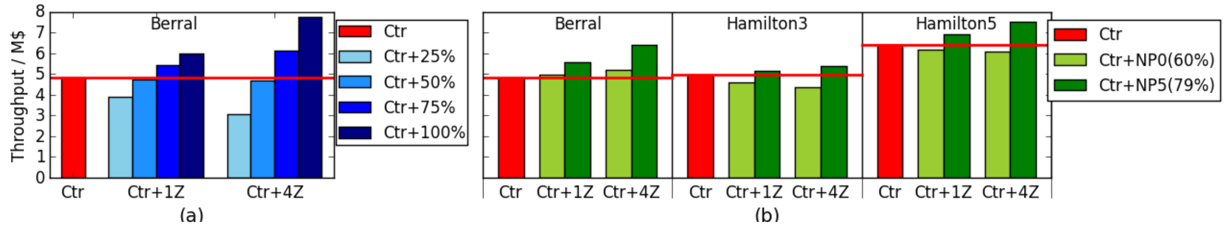


Fig. 19. (a) Throughput/M$ for periodic resources, varying duty factor and scale. (b) Throughput/M$ for irregular resources, varying SP model and scale.

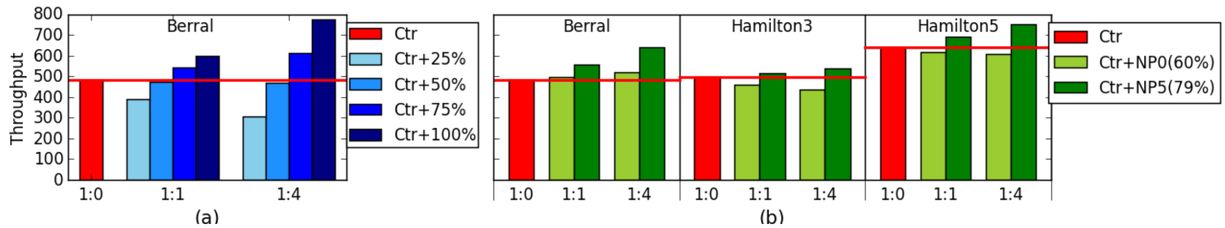


Fig. 20. (a) Throughput for periodic resources, and (b) Throughput for irregular resources, with fixed budget $100M of TCO. X-axis is the ratio of traditional datacenter MW to ZCCloud datacenter MW.

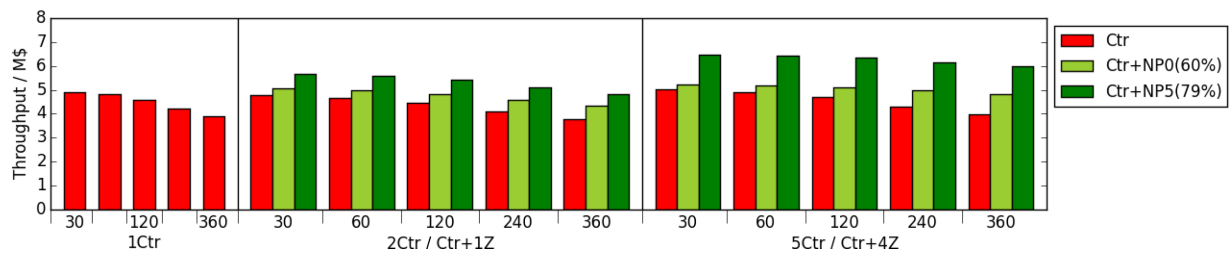


Fig. 21. Throughput/M$ vs. System scale vs. Power price (based on Berral's model). System scale increases from left to right. Power price from $30/MWh to $360/MWh.

lowers system capability. We explore the cost-performance of ZCCloud systems, for both periodic resources and irregular resources by combining productivity results from Section 4 with TCO models. Throughout, we use the metric *Throughput/Million $*, which is work per unit operating cost.

Periodic Intermittent Resources We first explore periodic resources, varying duty factor from 25% to 100%, and compare them to the baseline datacenter (Ctr). The results in Figure 19(a) show that as duty factor increases, cost-performance improves. At around 50% duty factor, the ZCCloud systems achieve comparable cost-effectiveness. As duty factor increases further, the ZCCloud systems become more cost-effective than traditional datacenters. Since many supercomputing projects have a fixed budget, and explore the most computationally powerful system design, we also compare the achieved throughput within a $100M budget in Figure 20 that confirms our conclusions.

Irregular Intermittent Resources For irregular resources, we consider NetPrice (NP) models as the LMP duty factors are uncompetitive. In Figure 19(b), we first consider Berral's model. Both NP0 and NP5 deliver duty factors >50%, and for NP0 the result is comparable or even slightly more cost-effective. NP5 has high duty factor of 79% and achieves 16% and 34% cost-effective advantage for Ctr+1Z and Ctr+4Z respectively. Then, for Hamilton3 and Hamilton5, the NP0-based ZCCloud systems are not quite competitive. This is because the lower TCO ratio of ZCCloud to Ctr resources in the Hamilton's models cannot compensate for intermittence losses. But for NP5, ZCCloud achieves superior cost performance, ranging 3-34% across three models.

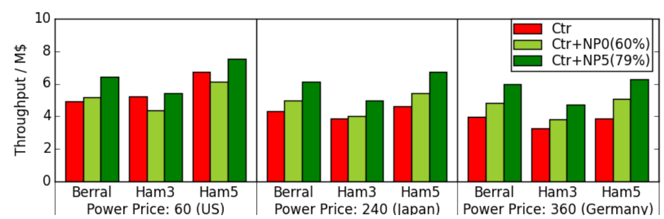


Fig. 22. Three TCO models for various power prices (system scale: Ctr+4Z).

In conclusion, expanding with ZCCloud intermittent resources with high duty factors can be more cost-effective than traditional resources. In the rest of this section, we explore varying power and hardware prices.

6.1 Cost-Performance vs. Power Price

Using Berral's model, Figure 21 compares cost-performance for small to large ZCCloud systems across a range of power price scenarios. Beginning with the red bars at the left, we can see that as the power price increases, system cost-effectiveness (throughput/M$) decreases. The next group shows that across the entire power price range, Ctr+1Z systems are more cost-effective than traditional, and the advantage grows as power price increases. This is because ZCCloud gets most of its power from stranded power. For NP5, the advantage is >28% at $360/MWh (i.e., Germany or Denmark). At the right side, we see that for larger systems the advantage also grows; the throughput/M$ of Ctr+4Z(NP5) is 28-50% higher than 5Ctr even at low power prices ($30-$60/MWh).

To compare across TCO models, we focus on US, Japan and Germany power prices (see Figure 22). With a typical US power price, under Hamilton's models, NP0-based systems are less cost-effective, but with NP5 they can be slightly better. However with more expensive electricity, the situation changes dramatically. When power price=$360/MWh, for Hamilton's models that are more power cost heavy, ZCCloud achieves 16-63% better cost-performance.

In conclusion, higher power prices make ZCCloud intermittent resources more cost-effective than traditional datacenter resources. NP0-based systems can sometimes be comparable, but NP5-based system give benefit across the board, and these benefits increase with higher power prices. Interestingly, Hamilton's models show even larger ZCCloud cost-performance benefits than Berral at high power prices. In high price power countries such as Japan and Germany, ZCCloud systems have a major cost-performance advantage, up to 63%. In more moderate-price regions the benefit can be 5-30%, depending on system size and TCO model.

6.2 Cost-Performance vs. Hardware Price

Using Berral's model, in Figure 23 we compare cost-performance across a wide range of hardware price. The results show a consistent trend: as hardware price increases throughput/M$ decreases. Because system cost-performance is more sensitive to compute hardware price, this drop is steeper than with power price. First, we consider lower hardware prices of 0.5x and 0.25x; such scenarios increase the advantage of ZCCloud dramatically. For example, at 0.25x, Ctr+1Z for NP5 achieves 41% better cost-performance than 2Ctr. For larger systems the advantage grows, reaching 91% for Ctr+4Z(NP5) vs. 5Ctr. Second, if hardware price increases, ZCCloud's benefits decrease, due to ZCCloud's poorer utilization. At 1.5x hardware price, ZCCloud systems are comparable with Ctr resources.

Considering all three TCO models in Figure 24, we focus on the extreme cheap and expensive hardware scenarios. With 0.25x compute hardware price, both Hamilton3 and Hamilton5 indicate large benefits for ZCCloud cost-efficiency, ranging 14-54%. For the high hardware price scenario, 1.5x, only the NP5-based ZCCloud can achieve a cost efficiency advantage.

In conclusion, lower hardware price increases the advantage of ZCCloud resources over traditional datacenter resources. When hardware price goes down due to the end of feature scaling, for Berral's model ZCCloud systems can be up to 91% more cost-effective. Hamilton's model is less favorable, but still achieves benefits as large as 54%.

7 ZCCloud at Extreme Scale

With the rapid penetration of computing into every facet of the economy, the power consumption of ICT generally and datacenters specifically is growing rapidly [16]. Scaling up represents major challenges, which requires expanding physical infrastructures, grabbing more power from the power grid, etc. In this section, we further explore the "post exascale scenario" and compare our ZCCloud approach with traditional scaling.

Supercomputing has set the goal of achieving exascale performance in 2020 (or 2022 in most recent US plan); the associated growth of power and required physical infrastructure are daunting. We consider two approaches: (Scenario A) traditional datacenters and, (Scenario B) ZCCloud - adding stranded power systems to complement a 4MW base system.

To create a baseline, we document performance and power for leading US Department of Energy systems: 2012 [60] and 2017 [21], and project based on geometric growth for 2022, 2027, and 2032 in Table 4. The resulting power numbers are daunting, despite projecting a continued 7x energy/op improvement every 5 years. Note that our model is quite conservative compared to Horst Simon's, former director of LBL NERSC and Deputy Director of LBL's empirical scaling model [78]. As described in Section 2.3, we expect nearly 200 MW of stranded power in a single ISO, suggesting perhaps 30x more (6 Gigawatts) may be available today, and growth in the future due to rising renewable penetration.

Figure 25 applies Berral's TCO model to extreme scaling, using a constant cost per unit for IT hardware, datacenter, and power. Specifically, TCO for the 39MW, 116MW and 232MW traditional systems rise to $430M, $1,300M, and $2,550M/year respectively. ZCCloud approach mitigates both power and datacenter elements of TCO growth, reducing TCO by 41% at 39MW and more than 45% at 232MW. While this estimate may overshoot hardware price (constant HW cost/MW), as demonstrated in Section 5.4, adjustments in this direction only increase the advantage of ZCCloud.

Applying Berral's model, and scaling with the rising power requirements, we consider cost-effectiveness in achieving "high peak performance", peak petaflops/million-dollars TCO (see Figure 26). It is a reflection of the peak capacity available for flexible computing (e.g., optimization, analytics). Our results indicate that the ZCCloud can increase the peak capability feasible at a given budget by as much as 80%.

For completness, we also consider cost-effectiveness on a throughput basis. Combining the productivity of ZCCloud hardware at duty factors feasible on stranded power (79%) with the TCO scaling, we compute throughput in jobs/million-dollars (see Figure 27). Our results show that the ZCCloud approaches not only achieve higher peak, but can be up to 45% more cost-effective on a overall compute per dollar TCO basis used to evaluate commercial datacenters.

TABLE 4
Projected Performance & Power, Top DOE Supercomputing Systems

Year	2012	2017	2022	2027	2032
Peak PF	10 [60]	200 [21]	4K	80K	?
MW	4 [60]	13 [21]	39	116	232
Horst Simon Model [78]					
Peak GF/kW	2.2K	4.2K	6.2K	8.2K	10.2K
MW	2	48	645	9.8K	?

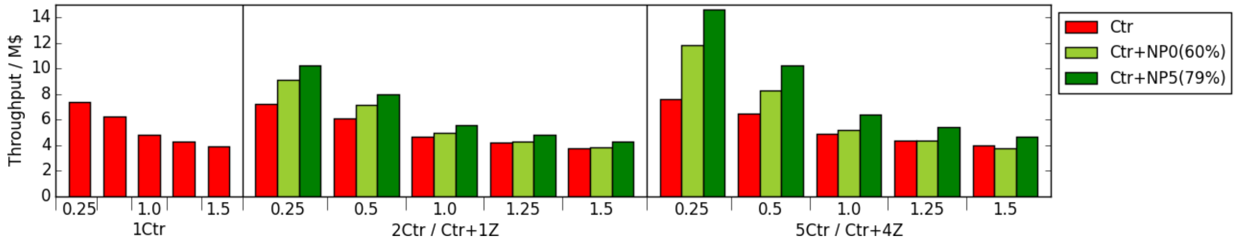


Fig. 23. Throughput/M$ vs. System scale vs. Compute HW price (based on Berral's model). System scale increases from left to right. Compute HW price ranges from 0.25x to 1.5x.

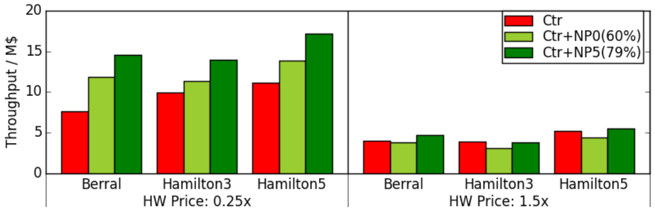


Fig. 24. Three TCO models for 0.25x and 1.5x compute HW price (system scale: Ctr+4Z).

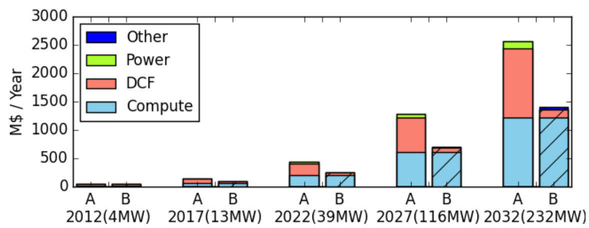


Fig. 25. TCO breakdown for Traditional (A) vs. ZCCloud (B).

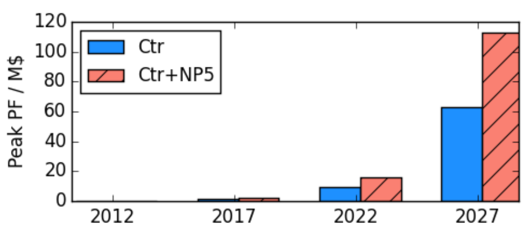


Fig. 26. Peak PFLOPS/M$ for traditional vs. ZCCloud, scaling TCO.

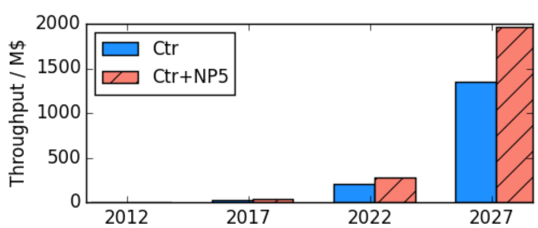


Fig. 27. Overall cost-competitiveness for theoretical throughput: traditional vs. ZCCloud, scaling TCO.

For brevity, we only consider extreme scaling with Berral's cost model, but our extensive study in Section 5 suggests that because the Hamilton model reflects power as a larger fraction of cost that we would see comparable benefits with the Hamilton3 (3-year) model, and even greater benefits with the Hamilton5 (5-year) model.

In conclusion, by exploiting stranded power, ZCCloud reduces the total cost of extreme-scale computing by up to 45%, improves the peak performance by 80% at a fixed budget, and is 45% more cost-effective.

8 Discussion and Related Work

Numerous researchers have studied integrating renewables into datacenters and maximizing green power usage [30], [31], [32], [33], [34], [35], [36]. More ambitious efforts even "follow the renewables" [51] distributing computing resources geographically to increase green power usage. ZCCloud differs from these efforts, using no grid power (brown or otherwise), exploiting only un-economic stranded power at the point of generation, and eliminating the cost of on-site renewable generation. Moreover, ZCCloud is a dispatchable load [1], [14] that may even benefit grid stability.

While power consumption limits the scaling of supercomputing, improving data-center energy-efficiency has become an area of intensive research. For example, underprovisioning reduces grid power supply to the average compute power demand instead of peak [79]. Dynamic power management selects power source based on the load [80], and energy-aware schedulers migrate jobs and shift peak load to achieve higher efficiency [36]. These approaches achieve higher power efficiency at some trade in system performance and quality of service, such as lowering CPU frequency and deferring jobs. In contrast, ZCCloud is designed as an extension to supercomputing systems to scale up capability and performance.

Distributed systems researchers have long sought to exploit intermittent computing resources, notably Peer-to-peer [38], desktop grid [39], [40], and Condor [81], where computing resources are generally single machines, loosely-coupled, and volatile. Cloud providers have also operated revocable computing services. Amazon's Spot Instances [42] enable cloud users to bid on EC2 instances with a lower cost, and revoke resource when its price exceeds the user's bid price. Google Compute Engine (GCE) provides analogous capability in a preemptible VM instance [43].

These resources have the similar fundamental unreliablity to those we consider. Remarkably, they have been widely used, and successfully applied to very large scale applications, such as Amazon Elastic MapReduce [44], Grid Computing [45], Fermilab's HEP Cloud project [46], Globus Genomics [47], etc. The recently introduced serverless computing services (Amazon's Lambda [48], Google Cloud Functions [49], and IBM's OpenWhisk [50]) are also potential matches to intermittent computing resources. In addition, there are a bunch of system studies that exploit intelligent replication, checkpointing, migration, and other techniques to make volatile resources more useful for cloud services requiring continuous availability [41], [82], [83], [84], [85], [86], [87], [88], [89]. We believe these efforts may have the potential to create high-value cloud services based on volatile ZCCloud resources.

Researchers have proposed migrating workloads around the world to reduce electricity cost at Akamai [90], and "follow-the-sun" techniques to increase the fraction of renewable energy use. Later techniques combine optimization for renewable use while meeting response time requirements [51]. All of these techniques suffer from TCO's dominated by the capital costs for compute hardware and physical infrastructure, and thus migratory techniques that produce low hardware resource utilization are generally not cost-effective. Our work on ZCCloud assesses achievable duty factors and overall costs, showing the ZCCloud can be more cost-effective than a single traditional supercomputing center facility.

Recently, several ISO's have initiated experimental energy storage programs [91]. Because storage economics are challenging, they focus on high-leverage applications – reducing peaks and frequency regulation. For example, California's statewide 5-year targets are 1.3GWx2h = 2.6GWh. In comparison, covering stranded power outages for a single ZCCloud unit might require 10MWx500h = 5GWh. At current prices, this energy storage cost is prohibitive ($100's M). We have a separate study [29] that details the limitations of using energy storage to compensate for large-scale intermittence. A parametric study which looks at how a partial solution of energy storage affects resource productivity and TCO is interesting, but beyond the scope of this paper.

Optimizing datacenter location cannot reduce stranded power. Because a steady load at a given location cannot resolve the dynamic problem of oversupply, transmission congestion, ramping constraints, or market dynamics that give rise to stranded power. Our study in IEEE Sustainable Energy [14] addresses this question directly, showing the limits of adding steady loads, even with extensive effort to pick optimal locations for the grids benefit (which of course cloud companies do NOT do today).

Our work builds on cost-model research for traditional and green datacenters. Specifically proposed frameworks to model and optimize cost/performance of a network of datacenters, considering capex and opex, and regional variations in electricity, networking, and real estate [63], [64].

9 Summary and Future Work

We explore cost-effectiveness of a novel approach to scaling supercomputing with stranded power. ZCCloud uses unreliable, stranded power by placing computing resources on generation side of the grid, reducing costs for power, power distribution, physical facilities, and a range of other costs.

Our study of ZCCloud TCO and cost-performance shows that compared to traditional datacenter resources, stranded-power based approaches can reduce total cost and increase cost-effectiveness in today's environment. Even without considering environment (reduced carbon emissions) and power-grid benefits (increased grid flexibility [14]), they can be cost-effective. This benefit is robust across TCO models and increases with scaleup of intermittent computing resources. Exploration of regional/future power prices and future compute hardware prices shows that ZCCloud's advantage in cost-performance grows with higher power price and cheaper compute hardware. The same trends are supported by all three TCO models. ZCCloud can be 43-63% more cost-effective at highest power price typical for Germany or Denmark, and 41-91% more cost-effective with 0.25x compute hardware price. Applying the TCO models to study future extreme scale supercomputing systems indicates the promise of the approach, enabling peak capability increases of 80% and cost-effectiveness increases of 45%.

Beyond analysis, it is valuable to build and operate ZCCloud prototypes to prove out the costs and effectiveness on real workloads. The first generation is a 9-node, 108-core ZCCloud prototype that executes Open Science Grid (OSG) [92] workload with availability driven by historical MISO stranded power data. Our second generation planned prototype is operational in fall 2016, and includes 78 nodes (nearly 2,000 cores), and will execute batch HPC workloads, OSG workloads, and intermittent cloud services. The next step is to design and build a physical prototype, and deploy it at a carefully selected wind farm with ample stranded power.

References

[1] A. A. Chien, R. Wolski, and F. Yang, "The zero-carbon cloud: High-value, dispatchable demand for renewable power generators," *The Electricity Journal*, pp. 110–118, 2015.

[2] ——, "Zero-carbon cloud: A volatile resource for high-performance computing," in *SHPC'15*. IEEE, 2015.

[3] F. Yang and A. A. Chien, "ZCCloud: Exploring Wasted Green Power for High-Performance Computing," in *IPDPS'16*, 2016.

[4] Intergovernmental Panel on Climate Change, "Climate change 2014: Synthesis report," 2014. [Online]. Available: http://www.ipcc.ch

[5] A. Gore, D. Guggenheim, L. David, L. Bender, S. Z. Burns, J. Skoll, L. Chilcott, B. Richman, J. Cassidy, and D. Swietlik, *An Inconvenient Truth*. Paramount, 2007.

[6] United Nations Framework Convention on Climate Change. (1997) Kyoto protocol. [Online]. Available: http://unfccc.int/kyoto_protocol/items/2830.php

[7] ——. (2015) Paris climate change conference. [Online]. Available: http://unfccc.int/meetings/paris_nov_2015/meeting/8926.php

[8] C. Megerian and J. Panzar, "Gov. Brown signs climate change bill to spur renewable energy, efficiency standards," Los Angeles Times, 9 2015.

[9] N. Y. S. E. P. Board, "The energy to lead: 2015 new york state energy plan," 2015, http://energyplan.ny.gov/Plans/2015.aspx.

[10] L. Bird, J. Cochran, and X. Wang, "Wind and Solar Energy Curtailment: Experience and Practices in the United States," NREL, Tech. Rep., March 2014.

[11] L. Bird, M. Milligan, and D. Lew, "Integrating Variable Renewable Energy: Challenges and Solutions," NREL, Tech. Rep., 2013.

[12] GWEC, "Global wind report: Annual market update," Global Wind Energy Council, Tech. Rep., 2016, documents curtailment around the world.

[13] E3, "Investigating a higher renewables portfolio standard in california: Executive summary," Report from Energy and Economics, Inc., Tech. Rep., January 2014.

[14] K. Kim, F. Yang, V. Zavala, and A. A. Chien, "Data centers as dispatchable loads to harness stranded power," *IEEE Transactions on Sustainable Energy*, 2016, dOI 10.1109/TSTE.2016.2593607.

[15] H. Dunn, "The carbon footprint of ICTs," *Global Information Society Watch*, 2010.

[16] Greenpeace, "Clicking clean: A guide to building a green internet, 2015 edition," 2015, http://www.greenpeace.org/usa/wp-content/uploads/legacy/Global/usa/planet3/PDFs/2015ClickingClean.pdf.

[17] J. Shalf, S. Dosanjh, and J. Morrison, "Exascale computing technology challenges," in *VECPAR 2010*, 2011, pp. 1–25.

[18] S. Borkar and A. A. Chien, "The future of microprocessors," *Commun. ACM*, vol. 54, May 2011.

[19] NUDT, "TIANHE-2," http://www.top500.org/system/177999, June 2015.

[20] J. Rath, "Blue Waters: Awesome Power, Awesome Efficiency," http://www.datacenterknowledge.com/archives/2010/06/24/blue-waters-awesome-power-awesome-efficiency/, June 2010, 15 MW power.

[21] ALCF, "Aurora," http://aurora.alcf.anl.gov/, 2015.

[22] OLCF, "Summit," https://www.olcf.ornl.gov/summit/, 2015.

[23] N. Bates, G. Ghatikar, G. Abdulla, G. A. Koenig, S. Bhalachandra, M. Sheikhalishahi, T. Patki, B. Rountree, and S. Poole, "The electrical grid and supercomputing center: An investigative analysis of emerging opportunities and challenges," in *3rd D-A-CH Conference on Energy Informatics*. Springer, 2014.

[24] O. Kimura and K.-i. Nishio, "Saving Electricity in a Hurry: A Japanese Experience after the Great East Japan Earthquake in 2011," *ACEEE Summer Study on Energy Efficiency in Industry*, 2013.

[25] NCSA, "Blue waters," https://en.wikipedia.org/wiki/Blue_Waters, June 2010.

[26] NERSC, "Computational and Theory Research Facility," https://www.nersc.gov/assets/Uploads/CRT-for-NUG.pdf, February 2015.

[27] D. Terdiman, "Is Google building a hulking floating data center in SF Bay?" *CNET*, 2013, http://www.cnet.com/news/is-google-building-a-hulking-floating-data-center-in-sf-bay/.

[28] J. Markoff, "Microsoft Plumbs Oceans Depths to Test Underwater Data Center," *The New York Times*, 2016.

[29] A. A. Chien, F. Yang, and C. Zhang, "Characterizing curtailed and uneconomic renewable power in the mid-continent independent system operator," arxiv.org, Tech. Rep., December 2016, available from .

[30] B. Aksanli, J. Venkatesh, L. Zhang, and T. Rosing, "Utilizing green energy prediction to schedule mixed batch and service jobs in data centers," *ACM SIGOPS Operating Systems Review*, vol. 45, no. 3, pp. 53–57, 2012.

[31] A. Krioukov, C. Goebel, S. Alspaugh, Y. Chen, D. E. Culler, and R. H. Katz, "Integrating renewable energy using data analytics systems: Challenges and opportunities." *IEEE Data Eng. Bull.*, vol. 34, no. 1, pp. 3–11, 2011.

[32] Z. Liu, Y. Chen, C. Bash, A. Wierman, D. Gmach, Z. Wang, M. Marwah, and C. Hyser, "Renewable and cooling aware workload management for sustainable data centers," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, no. 1. ACM, 2012, pp. 175–186.

[33] R. Sharma, T. Christian, M. Arlitt, C. Bash, and C. Patel, "Design of a farm waste-drive supply side infrastructure for data centers," in *ASME Energy Sustainability*, 2010.

[34] Í. Goiri, W. Katsak, K. Le, T. D. Nguyen, and R. Bianchini, "Parasol and greenswitch: Managing datacenters powered by renewable energy," in *ACM SIGARCH Computer Architecture News*. ACM, 2013, pp. 51–64.

[35] C. Li, A. Qouneh, and T. Li, "iSwitch: coordinating and optimizing renewable energy powered server clusters," in *ISCA'12*. IEEE, 2012.

[36] M. E. Haque, Í. Goiri, R. Bianchini, and T. D. Nguyen, "Greenpar: Scheduling parallel high performance applications in green datacenters," 2015.

[37] G. Cook, T. Dowdall, D. Pomerantz, and Y. Wang, "Clicking clean: how companies are creating the green internet," *Greenpeace Inc., Washington, DC*, 2014.

[38] A. Oram, Ed., *Peer-to-Peer: Harnessing the Power of Disruptive Technologies*. O'Reilly, 2001.

[39] A. Chien, B. Calder, S. Elbert, and K. Bhatia, "Entropia: architecture and performance of an enterprise desktop grid system," *JPDC*, pp. 597–610, 2003.

[40] D. P. Anderson, J. Cobb, E. Korpela, M. Lebofsky, and D. Werthimer, "SETI@home: an experiment in public-resource computing," *Communications of the ACM*, pp. 56–61, 2002.

[41] M. J. Litzkow, M. Livny, and M. W. Mutka, "Condor-a hunter of idle workstations," in *Distributed Computing Systems, 1988., 8th International Conference on*. IEEE, 1988, pp. 104–111.

[42] Amazon, "Spot instance," https://aws.amazon.com/ec2/spot/.

[43] Google, "Google compute engine: Creating a preemptible vm instance," https://cloud.google.com/compute/docs/instances/preemptible.

[44] Amazon, "Amazon elastic mapreduce (emr)," https://aws.amazon.com/emr/.

[45] ——, "Scientific computing using spot instances," https://aws.amazon.com/ec2/spot/spot-and-science/.

[46] B. Holzman, L. A. Bauerdick, B. Bockelman, D. Dykstra, I. Fisk, S. Fuess, G. Garzoglio, M. Girone, O. Gutsche, D. Hufnagel *et al.*, "Hepcloud, a new paradigm for hep facilities: Cms amazon web services investigation," *Computing and Software for Big Science*, vol. 1, no. 1, p. 1, 2017.

[47] R. K. Madduri, D. Sulakhe, L. Lacinski, B. Liu, A. Rodriguez, K. Chard, U. J. Dave, and I. T. Foster, "Experiences building globus genomics: a next-generation sequencing analysis service using galaxy, globus, and amazon web services," *Concurrency and Computation: Practice and Experience*, vol. 26, no. 13, pp. 2266–2279, 2014.

[48] Amazon, "Aws lambda," https://aws.amazon.com/lambda/.

[49] Google, "Google cloud functions," https://cloud.google.com/functions/.

[50] I. Bluemix, "Ibm cloud functions," https://www.ibm.com/cloud-computing/bluemix/openwhisk.

[51] J. L. Berral, Í. Goiri, T. D. Nguyen, R. Gavalda, J. Torres, and R. Bianchini, "Building green cloud services at low cost," in *ICDCS'14*. IEEE, 2014, pp. 449–460.

[52] J. Hamilton, "Overall data center costs," http://perspectives.mvdirona.com/2010/09/overall-data-center-costs/, 2010.

[53] Z. Shahan, "Renewable Energy = 13.4% of US Electricity Generation in 2014," http://cleantechnica.com/2015/03/10/renewable-energy-13-4-of-us-electricity-generation-in-2014-exclusive/.

[54] CPUC, "California Public Utilities Commission (CPUC)," http://www.cpuc.ca.gov/PUC/energy/Renewables.

[55] D. W. Program, "Wind vision: A new era for wind power in the united states," DOE National Renewable Energy Laboratory, http://energy.gov/eere/wind/wind-vision, Tech. Rep., May 2015.

[56] D. Lew, L. Bird, M. Milligan, B. Speer, X. Wang, E. M. Carlini, A. Estanqueiro, D. Flynn, E. Gomez-Lazaro, N. Menemenlis, A. Orths, I. Pineda, J. C. Smith, L. Soder, P. Sorensen, A. Altiparmakis, and Y. Yoh, "Wind and Solar Curtailment," in *Workshop on Large-Scale Integration of Wind Power Into Power Systems*, 10 2013.

[57] MISO, "Mid-continent ISO," https://www.misoenergy.org/.

[58] C. Krauss and D. Cardwell. (2015) A texas utility offers a nighttime special: Free electricity. New York Times. [Online]. Available: http://nyti.ms/1kDA8Iu

[59] Top500, "Top500 supercomputers," 2015, http://www.top500.org/.

[60] ALCF, "Mira," https://www.alcf.anl.gov/mira.

[61] L. Barroso, J. Clidaras, and U. Holzle, *The datacenter as a computer, 2nd Edition*. Morgan and Claypool, 2013.

[62] U. Holzle and L. Barroso, *The datacenter as a computer*. Morgan and Claypool, 2009.

[63] I. Goiri, K. Le, J. Guitart, J. Torres, and R. Bianchini, "Intelligent placement of datacenters for internet services," in *ICDCS'11*. IEEE, 2011, pp. 131–142.

[64] X. Wang, D. K. Chandrashekhara, M. Haque, I. Goiri, R. Bianchini, and T. Nguyen, "Grid-aware placement of datacenters and wind farms," in *IGSC'15*. IEEE, 2015.

[65] J. Park, "Open compute project: Data center v1.0," Open Compute Project, Tech. Rep., 2011, http://www.opencompute.org/.

[66] D. Bouley, "Estimating a data centers electrical carbon footprint," *Schneider Electric White Paper Library*, 2011, http://www.apc.com/salestools/DBOY-7EVHLH/DBOY-7EVHLH_R0_EN.pdf.

[67] ODCA, "Open Data Center Alliance Usage Model: Carbon Footprint and Energy Efficiency Rev. 2.0," Open Data Center Alliance, Tech. Rep., 2014, http://www.opendatacenteralliance.org/docs/Carbon-Footprint-and-Energy-Efficiency-Rev2.0.pdf.

[68] ALCF, "Cobalt scheduler," https://www.alcf.anl.gov/cobalt-scheduler.

[69] EIA, "EIA electricity price," http://www.eia.gov/electricity/, 2015.

[70] K. Fellman and T. Rideout, "Cost Estimates on Connecting Anchor Institutions to Fiber," The National Association of Telecommunications Officers and Advisors (NATOA), Tech. Rep., 2009, https://www.natoa.org/documents/NATOA%20Comments%20on%20NBP%2

[71] CTC, "A Model for Understanding the Cost to Connect Schools and Libraries with Fiber Optics," CTC Technology & Energy, Tech. Rep., 2014, http://www.ctcnet.us/wp-

content/uploads/2014/10/Connecting-Schools-and-Libraries-20141017.pdf.
[72] Intel, “Intel Enterprise SSD,” 2016, http://www.intel.com/content/www/us/en/solid-state-drives/solid-state-drives-ssd.html.
[73] Tesla, “Tesla powerwall,” 2016, https://en.wikipedia.org/wiki/Tesla_Powerwall.
[74] ALCF, “Argonne Leadership Computing Facility,” http://www.alcf.anl.gov/.
[75] O. Energy, “Average electricity prices around the world,” https://www.ovoenergy.com/guides/energy-guides/average-electricity-prices-kwh.html.
[76] S. Barbara, “No Moore?” *The Economist*, November 2013, world Print Edition.
[77] D. Clark, “Intel’s sway drives up server-chip prices,” *Wall Street Journal*, 2014.
[78] T. Trader, “Horst Simon on the HPC Slowdown,” http://www.hpcwire.com/2015/02/13/horst-simon-hpc-slowdown/, 2015.
[79] X. Zhou, Q. Cao, H. Jiang, and C. Xie, “Underprovisioning the grid power infrastructure for green datacenters,” in *ICS’15*. ACM, 2015, pp. 229–240.
[80] L. Liu, C. Li, H. Sun, Y. Hu, J. Gu, T. Li, J. Xin, and N. Zheng, “HEB: deploying and managing hybrid energy buffers for improving datacenter efficiency and economy,” in *ISCA’15*, 2015.
[81] HTCondor, “HTCondor,” http://research.cs.wisc.edu/htcondor/.
[82] X. He, P. Shenoy, R. Sitaraman, and D. Irwin, “Cutting the cost of hosting online services using cloud spot markets,” *HPDC’15*, 2015.
[83] W. Guo, K. Chen, Y. Wu, and W. Zheng, “Bidding for highly available services with low price in spot instance market,” *HPDC’15*, 2015.
[84] T. Trader, “Cycle computing orchestrates cancer research on google cloud,” Tech. Rep., 2015, http://www.hpcwire.com/2015/09/10/cycle-computing-orchestrates-cancer-research-on-google-cloud/.
[85] A. Shokrollahi, “Raptor codes,” *IEEE transactions on information theory*, vol. 52, no. 6, pp. 2551–2567, 2006.
[86] S. Yi, A. Andrzejak, and D. Kondo, “Monetary cost-aware checkpointing and migration on amazon cloud spot instances,” *IEEE Transactions on Services Computing*, vol. 5, no. 4, pp. 512–524, 2012.
[87] P. Sharma, S. Lee, T. Guo, D. Irwin, and P. Shenoy, “Spotcheck: Designing a derivative iaas cloud on the spot market,” in *Proceedings of the Tenth European Conference on Computer Systems*. ACM, 2015, p. 16.
[88] G. DeCandia, D. Hastorun, M. Jampani, G. Kakulapati, A. Lakshman, A. Pilchin, S. Sivasubramanian, P. Vosshall, and W. Vogels, “Dynamo: amazon’s highly available key-value store,” *ACM SIGOPS operating systems review*, vol. 41, no. 6, pp. 205–220, 2007.
[89] C. Huang, H. Simitci, Y. Xu, A. Ogus, B. Calder, P. Gopalan, J. Li, S. Yekhanin *et al.*, “Erasure coding in windows azure storage.” in *Usenix annual technical conference*. Boston, MA, 2012, pp. 15–26.
[90] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs, “Cutting the Electric Bill for Internet-scale Systems,” in *SIGCOMM’09*. ACM, 2009, pp. 123–134.
[91] CPUC, “Energy storage procurement framework and design program,” October 2013, http://www.cpuc.ca.gov/general.aspx?id=3462,
[92] OSG, “Open Science Grid,” http://www.opensciencegrid.org/, 2016.

APPENDIX
DETAILS OF TCO MODEL

The amortized costs in Table 3 are calculated using the following formulas:

$$C_{component} = \frac{r \cdot CapEx_{component}}{1-(1+r)^{-l}} \quad (5)$$

$$CapEx_{component} = Price_{component} \cdot Size_{component} \quad (6)$$

Where $CapEx_{component}$ is the capital expense of the component, r is the cost of capital, 3%, and l is the amortization period (in years). Table 5 includes full detail.

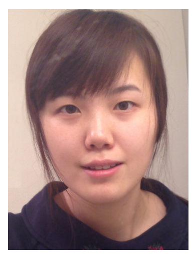
We present all elements from [63] in Table 6; our cost model omit those with contributions generally less that <1% including water, connection and land.

TABLE 5
Detailed Parameters of the Cost Model

Component	Price	Size	Amortization
Compute	$24M/MW	4MW	5 years
Network	$13k/mile	500 miles	10 years
SSD	$0.67/GB	2PB	5 years
Battery	$350/kWh	1MWh	5 years
Container	$5M/MW	4MW	12 years
Free Cooling	$700k/MW	4MW	10 years

TABLE 6
Full Cost Model from [63]

Elements	Description	Contribution
Components included in our simple model		
Servers	Computing hardware costs	44.9% - 56.9%
Building	Machine room, cooling and power infrastructures	23.6% - 34.5%
Power	Electricity costs	9.2% - 17.8%
Networking	Internal networking costs	0.7% - 1.1%
Connection	Connection to power grid and Internet backbone	0.1% - 4.7%
Components omitted in the simple model		
Water	Water used for cooling	0.1% - 1.2%
Land	Cost of real-estate	0.7% - 2.4%



Fan Yang is a fifth-year Ph.D. student in the Department of Computer Science, the University of Chicago. She joined the Large-Scale System Group at UChicago in 2013. She has broad research interests in cloud computing, database and storage system, parallel and distributed systems, data-intensive computing, large-scale system software, resilience and fault tolerance, green computing, and computer architecture. She currently focuses on exploiting green stranded power for large-scale computing, and scalable, highly-available data management on volatile computing resources. She received her B.Eng. in Computer Science at National University of Defense Technology, China, in 2012, and M.S. degree at the University of Chicago in 2015.

Andrew A. Chien is the William Eckhardt Professor in Computer Science and Director of the CERES Center for Unstoppable Computing at the University of Chicago and a Senior Computer Scientist at Argonne National Laboratory. Dr. Chien has broad interests in computer systems, including architecture, cloud and high-performance computing. He has published over 170 papers in these areas. From 2005 to 2010, Chien was Vice President of Research of Intel Corporation, and from 1998 to 2005, the SAIC Chair Professor in Computer Science and founder of the Center for Networked Systems at the University of California, San Diego. His research has been recognized for excellence with numerous awards, and also supported by the NSF, DARPA, DOE, ONR, NASA, and industry. Dr. Chien currently serves as co-chair of the editorial board of the Communications of the ACM. From 1990 to 1998, he was a Professor of Computer Science at University of Illinois. Dr. Chien earned B.S., M.S., and Ph.D. degrees at the Massachusetts Institute of Technology, and is a Fellow of the ACM, IEEE, and AAAS.