This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCC.2021.3132174, IEEE Transactions on Cloud Computing

Battery-Assisted Online Operation of Distributed Data Centers with Uncertain Workload and Electricity Prices

Jun Sun, *Student Member, IEEE,* Shibo Chen, *Member, IEEE,* Pengcheng You
Qinmin Yang, *Member, IEEE,* and Zaiyue Yang, *Member, IEEE*

Abstract—This paper investigates the online operation of distributed data centers equipped with energy battery. We aim to minimize their long-term operational cost by optimally distributing workload among data centers and operating energy battery. However, future spatio-temporally variant uncertainties in both workload and electricity prices have been the main impediment for a performance-guaranteed online data center operation strategy. To address this issue, we develop a fully distributed online algorithm that decouples workload distribution and battery operation across the network and time by introducing well-designed virtual queues for workload and batteries into the framework of Lyapunov optimization. Theoretically, an analytical gap between the long-term operational cost achieved by our algorithm and the theoretical optimum is provided to corroborate the desirable operation strategy. Extensive simulations using the real-world workload and electricity price data demonstrate the cost-delay tradeoff that our algorithm strikes and validate the theoretical results that we obtained.

Index Terms—Distributed data centers, battery, Lyapunov optimization.

1 Introduction

In recent years, cloud service providers have been driven to build data centers (DCs) by the continuously increasing demand for data processing worldwide. Huge energy consumption comes along with the prosperity of DCs, and the electricity bill has accounted for the majority of the total DC expenditure [1]. It would save cloud service providers millions of dollars to improve operational efficiency by even one percent [2], which has incentivized the industry to build and run green DCs in an energy-efficient way. There has been a large literature on improving energy efficiency of DCs from various perspectives, e.g., hardware and infrastructure [3], thermal management that reduces energy consumption of cooling systems [4], [5], dynamic resource allocation that mitigates energy demand [6], [7], and workload scheduling and alternative energy supply that cut down energy cost [8], [9].

However, most, if not all, of the above work [4]–[9] concentrates on one isolated DC, whose efficiency improvement has been approaching the bottleneck with all recent advances in theory and technology. Therefore, a network of distributed DCs that further exploits spatial variations of electricity prices and allows workload transfer has drawn increasing attention. The infrastructure of a distributed DC system, as shown in Fig.1, consists of mapping nodes (MNs) which receive and distribute workload, DCs which process workload, and the communication links between MNs and DCs for workload transfer. It is commonly modeled by a bipartite graph [10], [11]. In addition, energy supply for DCs usually comes from two sources: electricity purchased from the power grid and drawn from battery battery. Battery has become essential to DCs as backup to hedge against power failure [12]. However, a recent report [13] pointed out that these batteries are highly underutilized, whose ability to shift peak demand away from high-price periods has not been fully explored [14]–[16]. From above, distributed DCs are expected to take full advantage of (a) temporal and spatial price diversity, (b) battery and (c) deferability of workload to reduce operational cost.

Geographical load balancing has been well investigated for distributed DCs. [11] deals with the resource allocation problem of the distributed DCs from a stochastic learning and dual decomposition perspective, and develops a power procurement strategy. [17] integrates onsite renewable generation, electricity procurement from the grid and geographical workload balance to reduce energy cost of distributed DCs and increase renewable energy utilization. [18] and [19] jointly consider the day-ahead electricity bidding and the workload balance among DCs. [20], [21] coordinate workload processing to reduce electricity generation cost of a smart grid. Especially an online algorithm is developed using the predicted workload based on the receding horizon control (RHC) method in [22]. There are a lot more other perspectives of reducing operational cost for DCs, e.g., making use of battery [17], [23], participating in demand response programs [24], [25], and coordinating workload in both temporal and spatial domains [26], [27]. However, the existing

Jun Sun and Qinmin Yang are with college of Control Science and Engineering, the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou, China.
Shibo Chen and Zaiyue Yang are with with the Shenzhen Key Laboratory of Biomimetic Robotics and Intelligent Systems, Department of Mechanical and Energy Engineering, Southern University of Science and Technology, Shenzhen, 518055, China, and also with the Guangdong Provincial Key Laboratory of Human-Augmentation and Rehabilitation Robotics in Universities, Southern University of Science and Technology, Shenzhen, 518055.
Pengcheng You is with the Whiting School of Engineering, Johns Hopkins University, Baltimore, MD 21218, US.
This work was supported in part by the National Key Research and Development Program of China under Grant 2019YFB1705401; in part by the Natural Science Foundation of China under Grant 61873118; in part by the Science, Technology and Innovation Commission of Shenzhen Municipality under Grant ZDSYS20200811143601004 and Grant 20200925174707002; in part by Zhejiang Provincial Nature Science Foundation of China under Grant LZ21F030004; in part by Key-Area Research and Development Program of Guangdong Province under Grant No. 2018B010107002.
Corresponding author: Zaiyue Yang, email: yangzy3@sustech.edu.cn.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCC.2021.3132174, IEEE Transactions on Cloud Computing

literature usually only covers one or two of the advantages of distributed DCs discussed above and a comprehensive cost reduction model is missing. Besides, uncertainties of future electricity prices and workload in real-time operation are not well addressed. Common assumptions are made that either the distributions of stochastic variables are known or their accurate prediction in a future window is available. These prerequisites are difficult to realize in practice and therefore render the corresponding methods unlikely to implement.

It is noteworthy that there is another line of relevant studies in energy harvesting [28]–[30] which share similar model with the one considered in this paper. An energy packet network (EPN) model is established in [28], and the workload distribution among the workstations is optimized. Taking into account the insufficient supply of intermittent energy, [29] introduces a mathematical model of a cascaded multihop network where each node gathers energy through harvesting. A new product-form solution (PFS) is derived for the joint probability distribution of energy availability, and job queue length for an N-node tandem system. Building upon G-network queuing theory, [30] proposes to optimize the work and energy flow in an EPN, and attains the equilibrium probability distribution for both the backlog of work and the backlog of energy. The goal of these works is usually to minimize the response time of the jobs without considering the locational energy diversity.

In comparison, this paper jointly exploits all the aforementioned three advantages for operational cost reduction of distributed DCs, i.e., specify the routing of workload and determine the usage of battery to complement power supply from the power grid such that the long-term operational cost of distributed DCs is minimized. Since future electricity prices and workload are not known in advance, the problem is formulated as a stochastic network optimization problem [31]. Our fundamental idea is to take advantage of spatio-temporal variations of electricity prices, battery as energy buffer, and flexibility of workload so as to alleviate the impact of uncertainties in future workload and electricity prices. We build on Lyapunov optimization to design a fully distributed online algorithm that well manages future uncertainties and is readily implementable in practice.

Lyapunov optimization has been broadly studied in online stochastic network optimization [31]–[34]. In particular, in the context of DC operation, [35], [36] employ Lyapunov optimization to address price and workload uncertainties. A comprehensive operation including the workload admittance and dispatch, the scaling of processing frequencies and the consolidation of servers and network connections is jointly considered in [37], [38]. Both geographical workload distribution and server management are addressed in a hierarchical two timescale model in [10]. [36] develops a distributed method to coordinate tenants colocated in one DC. One major concern about Lyapunov optimization is that it mainly tackles average performance, e.g., trades off between average queue length and average cost in our problem. However, it may perform extremely poor under some circumstances, e.g., delay workload significantly in the case of low workload arrival rate [39]. There is no guarantee on worst-case performance. In our work we come up with a novel idea of designing virtual queues to additionally bound worst-case workload queuing delay.

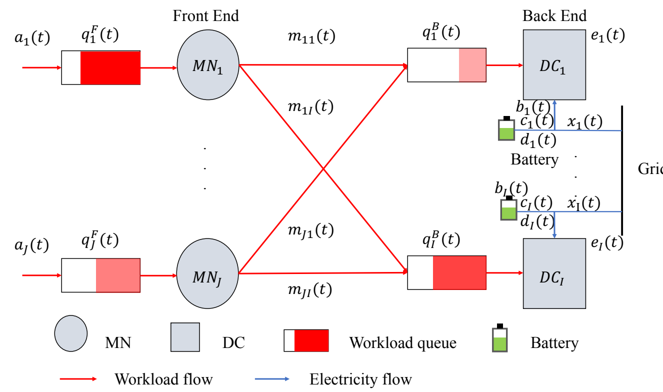


Fig. 1. The Infrastructure of Distributed Data centers

Our main contributions are summarized below.

- We investigate the battery-assisted online operation of distributed DCs in the presence of uncertain electricity prices and workload. Based on Lyapunov optimization, we develop a fully distributed online algorithm which bounds the gap between the long-term operational cost and its theoretical optimum.
- Novel virtual queues are designed in addition to real workload queues, such that the worst-case queuing delay can also be bounded in the online operation of distributed DCs. Therefore, the quality of service is enhanced.
- We further incorporate batteries to augment the flexibility of the distributed DC network. By constructing virtual queues for batteries and properly designing a Lyapunov function, we derive an explicit condition that guarantees the fulfillment of battery capacity constraints while realizing desirable long-term operational cost.
- The workload traces of Google DCs and the real-time electricity prices from the PJM ISO [40] [1] are employed to conduct the realistic simulations, which corroborate our theoretical analysis.

The remainder of the paper is organized as follows: Section II describes the mathematical models of the distributed DCs and formulates the long-term operational cost minimization problem. We propose our distributed online algorithm based on Lyapunov optimization and analyze its performance guarantee in Section III. Section IV presents the simulations on synthetic and real-world data. Finally, Section V concludes this paper.

2 Problem Formulation

2.1 Distributed Data Centers and Workload

A distributed DC system, as illustrated in Fig. 1, consists of a set of geographically distributed DCs, $\mathbb{I} = [1, 2, \cdots, I]$, a

1. PJM stands for the Pennsylvania, Jersey, Maryland Power Pool. At present, PJM market is the wholesale electricity market that operates an electric transmission system serving all or parts of Delaware, Illinois, Indiana, Kentucky, Maryland, Michigan, New Jersey, North Carolina, Ohio, Pennsylvania, Tennessee, Virginia, West Virginia, and the District of Columbia.

set of MNs, $\mathbb{J} = [1, 2, \cdots, J]$ and the communication links between the DCs and MNs. The workload from the users arrives at the MNs which distribute them among the DCs to process. In this way, the MNs and DCs play the roles of the front end and the back end, respectively. We denote the DCs connected with MN j and the MNs connected with DC i by Ω_j^F and Ω_i^B, respectively.

We only consider workload that is delay tolerant. Further assume the energy consumed to process each unit of workload is constant, then we are able to quantify workload by the amount of energy consumption. Workload is served in a First-In-First-Out (FIFO) fashion. That is, the workload that arrives at the MNs (DCs) earlier has higher priority to be transferred (processed). The DC operation includes workload distribution and power supply management, and the latter mainly involves electricity purchase and battery management. Workload distribution is to specify at each time slot the amount of workload each MN should transfer to each DC. Power supply management in each data center is to decide the amount of electricity to purchase from different local electricity markets for processing workload and the usage (charging/discharging) of battery.

We consider a slotted time horizon $\mathbb{T} = [1, 2, \cdots, T]$. At the beginning of time slot t, the amount of workload that arrives at MN j is denoted by $a_j(t)$, $a_j(t) \in [0, A_j]$, where A_j is the maximum amount of the workload that can enter MN j. After workload arrives at MNs, it will be dispatched to certain DCs to be processed. At time slot t, if MN j holds a workload queue with length $q_j^F(t)$, receives workload $a_j(t)$ and delivers workload $m_{ji}(t)$ to DC i, $i \in \Omega_j^F$, the remaining workload at $t+1$ is

$$q_j^F(t+1) = max\{q_j^F(t) + a_j(t) - \sum_{i \in \Omega_j^F} m_{ji}(t), 0\} \quad (1)$$

At time t, if DC i holds a workload queue with length $q_i^B(t)$, receives workload $m_{ji}(t)$ from MN j, $j \in \Omega_i^B$ and processes $e_i(t)$ amount of workload, the resultant queue at $t+1$ is

$$q_i^B(t+1) = max\{q_i^B(t) + \sum_{j \in \Omega_i^B} m_{ji}(t) - e_i(t), 0\} \quad (2)$$

Suppose the bandwidth cost for transferring workload is proportional to the amount transferred and denote the constant unit price between MN j and DC i by α_{ji}, which usually depends on the distance and communication condition of the link. The workload transferring cost of MN j is then $\sum_{i \in \Omega_j^F} \alpha_{ji} m_{ji}(t)$. Besides, the bandwidth of each link limits the maximum amount of transferable workload:

$$0 \leq m_{ji}(t) \leq M_{ji}, \forall j \in \mathbb{J}, \forall i \in \mathbb{I}, \forall t \in \mathbb{T} \quad (3)$$

2.2 Energy

The energy provision for DCs includes the electricity procured from the grid and the battery owned by DCs.

2.2.1 Electricity from the grid

Let $p_i(t)$ denote the real-time electricity price in the local electricity market where DC i is located. Denoted by $x_i(t)$ the amount of electricity purchased in time slot t by DC i, then the electricity bill is $p_i(t)x_i(t)$. Assume $p_i(t)$ is bounded in a compact set: $\underline{p_i} \leq p_i(t) \leq \overline{p_i}$.

2.2.2 Battery

We exploit the battery to improve operational efficiency of DCs. Let $b_i(t)$, $c_i(t)$ and $d_i(t)$ be the battery energy level, charged energy and discharged energy at time t in DC i, respectively. Then the following equation (4) characterizes the dynamics of the battery energy level:

$$b_i(t+1) = b_i(t) + \eta^c c_i(t) - \frac{1}{\eta^d} d_i(t) \quad (4)$$

where $\eta^c, \eta^d \in (0, 1)$ are the charging and discharging efficiencies, respectively.

We make a mild and fundamental assumption here:

Assumption 1. *The upper and lower bounds of electricity prices and the charging and discharging efficiencies of batteries satisfy*

$$\underline{p_i} < \overline{p_i} \cdot \eta^c \eta^d \quad (5)$$

Remark 1. *If* (5) *does not hold, namely,* $\overline{p_i}\eta^c\eta^d \leq \underline{p_i}$, *then* $p_i(t_1)\eta^c\eta^d y - p_i(t_2)y \leq (\overline{p_i}\eta^c\eta^d - \underline{p_i})y \leq 0, \forall y > 0$ *and* $\forall t_1, t_2 \in \mathbb{T}$, *which means the cost of electricity when charged is always no less than its revenue when discharged, causing negative profit. In this case, batteries do not help reduce the operational cost of DCs and will remain idle in any optimal operation. Therefore, Assumption 1 is made to highlight the benefit of operating batteries.*

We require that each battery i[2] should maintain at least a certain level of energy $\underline{b_i}$. Denote its capacity by $\overline{b_i}$, then $b_i(t)$ needs to satisfy

$$\underline{b_i} \leq b_i(t) \leq \overline{b_i}, \forall i \in \mathbb{I}, \forall t \in \mathbb{T} \quad (6)$$

Note that constraint (6) couples the operation of batteries over the whole time horizon, which brings a great challenge to the design of online algorithms. To decouple across time slots, a virtual queue method is designed, as we will explain in Subsection 3.2.

The cost of batteries is quantified by their usage. Let $c_i^{ON}(t)$ be a binary variable denoting the charging state of battery i in slot t. If $c_i^{ON}(t) = 1$, battery i is being charged; otherwise, it is not being charged. We define the corresponding discharging variable $d_i^{ON}(t)$ similarly. Let β represent the battery loss coefficient, then the cost of battery i is $\beta(c_i^{ON}(t) + d_i^{ON}(t))$. The variables, $c_i(t), d_i(t), c_i^{ON}(t), d_i^{ON}(t)$, are constrained as follows:

$$c_i^{ON}(t), d_i^{ON}(t) \in \{0, 1\}, \forall i \in \mathbb{I}, \forall t \in \mathbb{T} \quad (7)$$

$$0 \leq c_i(t) \leq c_i^{ON}(t) \cdot \overline{c_i}, \forall i \in \mathbb{I}, \forall t \in \mathbb{T} \quad (8)$$

$$0 \leq d_i(t) \leq d_i^{ON}(t) \cdot \overline{d_i}, \forall i \in \mathbb{I}, \forall t \in \mathbb{T} \quad (9)$$

$$0 \leq c_i^{ON}(t) + d_i^{ON}(t) \leq 1, \forall i \in \mathbb{I}, \forall t \in \mathbb{T} \quad (10)$$

(8) and (9) ensure that the charging and discharging rates are nonnegative and upper bounded by $\overline{c_i}$ and $\overline{d_i}$, respectively, and (10) indicates that no simultaneous charging and discharging are allowed.

Meanwhile, the energy consumption and provision in DC i should be balanced:

$$e_i(t) = x_i(t) + d_i(t) - c_i(t), \forall i \in \mathbb{I}, \forall t \in \mathbb{T} \quad (11)$$

2. We abuse i to index both DCs and batteries without ambiguity.

Obviously, $e_i(t)$ should be nonnegative and upper bounded due to the limited workload processing capability of DC i, i.e.,

$$0 \le e_i(t) \le E_i \tag{12}$$

For brevity, we collect the stochastic parameters, $a_j(t)$ and $p_i(t)$, in $\mathcal{S}(t), t \in \mathbb{T}$, and summarize the decision variables, $x_i(t), c_i^{ON}, d_i^{ON}(t), c_i(t), d_i(t)$ and $m_{ji}(t)$, in $\mathcal{X}(t)$. Note that we do not assume any prior knowledge of distributions for $\mathcal{S}(t)$, thus our method is quite general. The network should be stabilized under the online decision $\mathcal{X}(t)$ in the sense of stable and finite queues, i.e.,

$$\lim_{T\to\infty} \frac{1}{T} \sum_{t\in\mathbb{T}} \{ \sum_{j\in\mathbb{J}} \mathbb{E}[q_j^F(t)] + \sum_{i\in\mathbb{I}} \mathbb{E}[q_i^B(t)] \} < \infty \tag{13}$$

where $\mathbb{E}[\cdot]$ is the expectation taken with respect to the uncertain $\mathcal{S}(t)$.

2.3 Problem Formulation

Based on above model, the total cost of the distributed DC system for each time slot is composed of the electricity bill, battery loss and bandwidth cost:

$$f(t) = \sum_{i\in\mathbb{I}} [p_i(t)x_i(t) + \beta(c_i^{ON}(t) + d_i^{ON}(t)) + \sum_{j\in\Omega_i^B} \alpha_{ji} m_{ji}(t)]$$

The long-term operational cost minimization problem is then formulated as follows:

$$\begin{aligned} \textbf{P1}: \quad & \min_{\mathcal{X}(t)} \quad \lim_{T\to\infty} \frac{1}{T} \sum_{t\in\mathbb{T}} \mathbb{E}[f(t)] \\ & s.t. \quad (1)-(4), (6)-(13) \end{aligned}$$

Problem **P1** is difficult for three reasons: 1) **P1** is a large-scale mixed-integer linear programming (MILP) problem, which is computationally prohibitive to solve directly; 2) The future electricity prices and workload are not known *a prior*; 3) All time slots are coupled by the battery capacity constraint (6).

Before proceeding to design an efficacious online algorithm that provably works well for **P1**, we derive a relaxation of **P1** that is easier to characterize as a benchmark. First, we relax constraints (4) and (6) into the following equation:

$$\lim_{T\to\infty} \eta^c \frac{1}{T} \sum_{t\in\mathbb{T}} \mathbb{E}[c_i(t)] = \lim_{T\to\infty} \frac{1}{\eta^d} \frac{1}{T} \sum_{t\in\mathbb{T}} \mathbb{E}[d_i(t)], \forall i \in \mathbb{I} \tag{14}$$

(14) can be obtained by summing (4) over $t = 1, 2, \cdots, T$, dividing by T, and taking the expectation with T going to infinity. The initial and final battery energy levels, divided by T, vanish as $T \to \infty$.

It can be verified that (14) is a necessary but not sufficient condition of (4) and (6). Therefore, the following **P2** is a relaxation of **P1**.

$$\begin{aligned} \textbf{P2}: \quad & \min_{\mathcal{X}(t)} \quad \lim_{T\to\infty} \frac{1}{T} \sum_{t\in\mathbb{T}} \mathbb{E}[f(t)] \\ & s.t. \quad (1)-(3), (7)-(14) \end{aligned}$$

Although **P2** is a relaxed version of **P1**, it is still difficult to characterize the property of its solution. In order to provide a benchmark, we further assume that $\mathcal{S}(t)$ is from an identical and independent distribution (I.I.D.). Then following theorem characterizes the solution to **P2**. Note that the proofs of all the theorems in this paper can be found in the supplementary materials.

Theorem 1. *Given that the stochastic parameter $\mathcal{S}(t)$ follows I.I.D. over $\mathbb{T}$, there exists a stationary randomized decision $\hat{\mathcal{X}}(t)$, which requires only the current information $\mathcal{S}(t)$ and satisfies*

$$\eta^c \mathbb{E}[\hat{c}_i(t)] = \frac{1}{\eta^d} \mathbb{E}[\hat{d}_i(t)] \tag{15}$$

$$\mathbb{E}[\sum_{i\in\Omega_j^F} \hat{m}_{ji}(t)] = \mathbb{E}[a_j(t)] \tag{16}$$

$$\mathbb{E}[\sum_{j\in\Omega_i^B} \hat{m}_{ji}(t)] = \mathbb{E}[\hat{e}_i(t)] \tag{17}$$

$$\mathbb{E}[f(t)] = \hat{f} \tag{18}$$

*where $\hat{f}$ denotes the optimal objective value of **P2**.*

The proof of Theorem 1 follows directly from [41] and is omitted here for brevity. Equations (15)-(17) indicate the energy and workload balance: the amount of electricity drawn from a battery equals the amount injected into the battery; the amount of workload distributed by a MN equals the amount it receives; the amount of workload processed by a DC equals the amount it is assigned. However, it should be noted that the strategy $\hat{\mathcal{X}}(t)$ may be infeasible for **P1**, since the upper bound and lower bound of battery energy levels are ignored and may be violated. The solution to **P2** therefore low bounds that to **P1** in terms of the long-term operational cost. Denote by f^* the optimal objective value of **P1**, then obviously $\hat{f} \le f^*$. Therefore, we are able to benchmark against $\hat{f}$ instead of f^*. To be more specific, denoted by $\tilde{f}$ the average cost of one online algorithm, then the fact that $\tilde{f} - \hat{f}$ is bounded suffices to show $\tilde{f} - f^*$ is bounded.

Next we proceed to design an online solution to problem **P1** based on Lyapunov optimization.

3 Distributed Online Algorithm

3.1 Lyapunov Optimization preliminary

The Lyapunov optimization method is developed for long term stochastic dynamic network optimization, where the input and the parameters of the network are random variables, and the goal is to stabilize the network and meanwhile optimize the utility (minimize the average energy consumption [32], [42], minimize average delay [33], or maximize the average throughput [34]). In the Lyapunov optimization scheme, a Lyapunov function is defined to quantify the network potential, which is a measurement of the queues. Minimizing the upper bound of Lyapunov drift (defined as the increment of Lyapunov function, as shown in (24)) promises to stabilize the network, because it maintains the network potential below a certain level. When the potential moves above this level, the Lyapunov drift will be negative, which drives the network potential back. When it comes to stabilizing the network and meanwhile minimizing the operational cost of the network, the upper bound of the drift-plus-penalty (weighted sum of the drift and the cost) is minimized, which is a tradeoff between the potential level and the cost. Greedily minimizing the penalty-plus-drift results in a dynamic controller which only requires the current information of the network. For a more comprehensive introduction of Lyapunov optimization, please refer to [43], [44].

3.2 Construction of Virtual Queues and Lyapunov Function

It is not straightforward to directly apply the Lyapunov optimization method to problem **P1** for two reasons: 1) The constraint (6) is time-coupling, which restrains us from designing an online algorithm with performance guarantee; 2) Standard Lyapunov optimization bounds only the average workload queuing delay, but cannot deal with special cases where the workload arrival rate is low, thereby causing extremely large delay. To cope with these issues, virtual workload queues and virtual battery queues are introduced before defining the Lyapunov function for our problem.

In order to tackle the issue of potential large workload queuing delay, we define the virtual workload queues $z_j^F(t)$ and $z_i^B(t)$ associated with $q_j^F(t)$ and $q_i^B(t)$, respectively.

$$z_j^F(t+1) = max\{z_j^F(t) - \sum_{i\in\Omega_j^F} m_{ji}(t) + \varepsilon_j^F \mathbb{1}_{q_j^F(t)}, 0\} \quad (19)$$

$$z_i^B(t+1) = max\{z_i^B(t) - e_i(t) + \varepsilon_i^B \mathbb{1}_{q_i^B(t)>0}, 0\} \quad (20)$$

where ε_j^F and ε_i^B are constant weights, which we shall explain later. $\mathbb{1}_x$ is an indicator function with $\mathbb{1}_x = 1$ if $x > 0$ and $\mathbb{1}_x = 0$ otherwise. The initial virtual queue lengths $z_j^F(0)$ and $z_i^B(0)$ are 0. By definition, the virtual queue has a positive arrival rate of workload ε_j^F (ε_i^B) if the real queue in MN j (DC i) is not empty. As we will show later, virtual workload queues can ensure a bounded worst-case workload queuing delay, especially when the workload arrival rate is low.

To deal with the time-coupling constraint (6), we further introduce a virtual queue $B_i(t)$ for each battery:

$$B_i(t) = b_i(t) - \underline{b_i} - \eta^d V \overline{p_i} + \eta^c \overline{c_i} \quad (21)$$

where V is a constant to be designed. $B_i(t)$ therefore only depends on $b_i(t)$ and its dynamics are also similar:

$$B_i(t+1) = B_i(t) + \eta^c c_i(t) - \frac{1}{\eta^d} d_i(t) \quad (22)$$

Let $\Theta(t) = [q_j^F(t), z_j^F(t), q_i^B(t), z_i^B(t), B_i(t)]$ ($\forall j \in \mathbb{J}, i \in \mathbb{I}$) be a state vector collecting all the queue information in the network at time t. Now we define our specific Lyapunov function as

$$\begin{aligned} L(\Theta(t)) =& \frac{1}{2}\sum_{j\in\mathbb{J}}[q_j^F(t)^2 + z_j^F(t)^2] \\ &+ \frac{1}{2}\sum_{i\in\mathbb{I}}[q_i^B(t)^2 + z_i^B(t)^2 + rB_i(t)^2] \end{aligned} \quad (23)$$

where r is a scaling parameter of virtual battery queues. The Lyapunov function $L(\Theta(t))$ is a potential function that quantifies the overall queue lengths in the network. As we will show later, by carefully choosing the parameters V and r we are able to satisfy the constraint (6) while providing bounded performance guarantee.

Remark 2. *The intuition is that by Lyapunov optimization, the Lyapunov function $L(\Theta(t))$, as well as $B_i(t)$, is bounded, thus a carefully chosen (V, r) pair can guarantee that the real battery energy level $b_i(t)$, the inverse mapping from $B_i(t)$, is also bounded and always satisfies* (6). *By this means, we can ignore the time-coupling constraint and solve the local problem at each slot in an online manner.*

3.3 Distributed Online Algorithm

The Lyapunov drift is defined as the conditional expectation of the 1-slot increment of the Lyapunov function:

$$\Delta(t) = \mathbb{E}[L(\Theta(t+1)) - L(\Theta(t))|\Theta(t)] \quad (24)$$

By minimizing an upper bound on the Lyapunov drift, we figure out a one-step operation with approximately the maximum decrease or the minimum increase in the Lyapunov function (potential of queues), given the current system states $\Theta(t)$. This operation is basically a local optimum that drives the network to a stable equilibrium.

Further, the drift-plus-penalty is then defined as the weighted sum of the Lyapunov drift and the one-step operational cost given $\Theta(t)$:

$$\Delta(t) + V\mathbb{E}[f(t)|\Theta(t)]$$

where V functions as a weight between the long-term operational cost and the overall queuing delay. By minimizing the upper bound on the drift-plus-penalty, we strike a tradeoff between minimizing the long-term operational cost and maintaining the network stable. The following lemma gives an upper bound of the drift-plus-penalty.

Lemma 1. *The drift-plus-penalty is upper bounded by*

$$\begin{aligned} \Delta(t) + V\mathbb{E}[f(t)|\Theta(t)] \le& N_1 + rN_2 + V\mathbb{E}[f(t)|\Theta(t)] \\ &+ \sum_{j\in\mathbb{J}} q_j^F(t)\mathbb{E}[(a_j(t) - \sum_{i\ \in\Omega_j^F} m_{ji}(t))|\Theta(t)] \\ &- \sum_{j\in\mathbb{J}} z_j^F(t)\mathbb{E}[\sum_{i\ \in\Omega_j^F} m_{ji}(t)|\Theta(t)] \\ &+ \sum_{i\in\mathbb{I}} q_i^B(t)\mathbb{E}[(\sum_{j\ \in\Omega_i^F} m_{ji}(t) - e_i(t))|\Theta(t)] \\ &- \sum_{i\in\mathbb{I}} z_i^B(t)\mathbb{E}[(e_i(t)|\Theta(t)] \\ &+ \sum_{i\in\mathbb{I}} rB_i(t)\mathbb{E}[(\eta^c c_i(t) - \frac{1}{\eta^d} d_i(t))|\Theta(t)] \end{aligned} \quad (25)$$

where N_1 and N_2 are constant defined below:

$$\begin{aligned} N_1 =& \frac{1}{2}\sum_{j\in\mathbb{J}}[max\{A_j^2, {M_j^F}^2\} + max\{{\varepsilon_j^F}^2, {M_j^F}^2\}] \\ &+ \frac{1}{2}\sum_{i\in\mathbb{I}}[max\{(\sum_{j\in\Omega_i^B} M_{ji} + \bar{c}_i)^2, (\bar{x}_i + \bar{d}_i)^2\} \\ &+ max\{(\varepsilon_i^B + \bar{c}_i)^2, (\bar{x}_i + \bar{d}_i)^2\}] \\ &+ \sum_{j\in\mathbb{J}} Z_j^F \varepsilon_j^F + \sum_{i\in\mathbb{I}} Z_i^B \varepsilon_i^B \\ N_2 =& \frac{1}{2}\sum_{i\in\mathbb{I}} max\{(\eta^c \bar{c}_i)^2, (\frac{1}{\eta^d}\bar{d}_i)^2\} \end{aligned}$$

and $M_j^F = \sum_{i\in\Omega_j^F} M_{ji}$*,* $M_i^B = \sum_{j\in\Omega_i^B} M_{ji}$*;* Z_i^B *and* Z_j^F *are respectively the upper bounds of* $z_i^B(t)$ *and* z_j^F*.*

Proof. From (1), we have:
$q_j^F(t+1)^2 \le [q_j^F(t) + a_j(t) - \sum_{i\in\Omega_j^F} m_{ji}(t)]^2$, thus,

$$\frac{1}{2}[q_j^F(t+1)^2 - q_j^F(t)^2] \le \frac{1}{2}[(a_j(t) - \sum_{i\in\Omega_j^F} m_{ji}(t))^2 + 2q_j^F(t)(a_j(t) - \sum_{i\in\Omega_j^F} m_{ji}(t))] \le \frac{1}{2}max\{\overline{a}_j^2, {M_j^F}^2\} + q_j^F(t)(a_j(t) - \sum_{i\in\Omega_j^F} m_{ji}(t))$$

Similar bounds can be derived for other queues. (25) is then straightforward by summing them up, taking the conditional expectation and adding $V\mathbb{E}[f(t)|\Theta(t)]$ on both sides. □

Given current queue states $\Theta(t)$ and the observed stochastic parameters $\mathcal{S}(t)$ at time t, minimizing the right-hand side of (25) is equivalent to solving the following local problem:

$$\textbf{P3}: \min_{\mathcal{X}(t)} \quad \sum_{i\in\mathbb{I}} g_i(t) + \sum_{j\in\mathbb{J}}\sum_{i\in\Omega_j^F} g_{ji}(t)$$
$$s.t. \quad (3), (7)-(12)$$

where $g_i(t)$ and $g_{ij}(t)$ follows directly below from the right-hand side of (25) by dropping the constant terms and the expectation operator, since current $\mathcal{S}(t)$ has been realized.

$$g_i(t) = Vp_i(t)x_i(t) + (q_i^B(t) + z_i^B(t))(-x_i(t) - d_i(t) + c_i(t)) + rB_i(t)(\eta^c c_i(t) - \frac{1}{\eta^d}d_i(t)) + V\beta(c_i^{ON}(t) + d_i^{ON}(t))$$

$$g_{ji}(t) = (V\alpha_{ji} + q_i^B(t) - q_j^F(t) - z_j^F(t))m_{ji}(t)$$

Therefore, **P3** is equivalent to trading off between minimizing the operational cost and maintaining stable and finite workload queues. The network would be stabilized by implementing the solution of **P3**, since once the queues grow too long, more electricity will be provided to accelerate processing the workload, as implied by $g_i(t)$ and $g_{ji}(t)$. Note that the time-coupling constraint (6) is not involved in **P3**.

By this means, the original time-coupled problem can be solved in an online fashion. Next we show that problem **P3** can be further tackled in a distributed manner. Denote by $\mathcal{X}_1(t)$ and $\mathcal{X}_2(t)$ the sets that summarize all the local operation variables of DCs i, $x_i(t), c_i^{ON}(t), c_i(t), d_i^{ON}(t), d_i(t)$ and the network routing variables, $m_{ji}(t)$, respectively. The problem **P3** naturally decomposes into two subproblems: the local power management (**LPM**):

$$\textbf{LPM}: \min_{\mathcal{X}_1(t)} \sum_{i\in\mathbb{I}} g_i(t)$$
$$s.t. \quad (7)-(12)$$
$$e_i(t) \le q_i^B(t) + \sum_{j\in\Omega_i^B} m_{ji}(t), \forall i \in \mathbb{I}$$

Algorithm 1 Distributed Online Operation Strategy of Distributed DCs

Initialization: Set $t=0$, $z_j^F(t)=0, q_j^F(t)=0 (j\in\mathbb{J})$, $z_i^B(t)=0, q_i^B(t)=0, B_i(t) = -\underline{b_i} - \eta^d\overline{p_i}V + \eta^c\overline{c_i} (i\in\mathbb{I})$.
while ($t\le T$) **do**
  For each MN j: receive the queue information from all DCs i, $i\in\Omega_j^F$, and solve its **NWR** problem.
  For each DC i: solve its **LPM** problem.
  Update the queues according to (1), (2), (19), (20) and (22).
  Set $t = t+1$;
end while

and the network workload routing (**NWR**):

$$\textbf{NWR}: \min_{\mathcal{X}_2(t)} \sum_{i\in\mathbb{I}}\sum_{j\in\Omega_i^B} g_{ij}(t)$$
$$s.t. \quad (3),$$
$$\sum_{i\in\Omega_j^F} m_{ji}(t) \le q_j^F(t) + a_j(t), \forall j\in\mathbb{J}$$

Remark 3. *The advantages of Lyapunov optimization are threefold: 1) The introduction of virtual battery queues allows us to remove the time-coupling constraint* (6)*. Consequently, the problem reduces to dynamically minimizing the local drift-plus-penalty in an online fashion; 2) At each time slot, since current queues and stochastic parameters have been observed, the conditional expectations are also removed; 3)Problem* ***P3*** *can be decomposed into* ***LPM*** *and* ***NWR****, both of which can be solved in a distributed manner.*

3.4 Distributed Implementation and Computational Effort

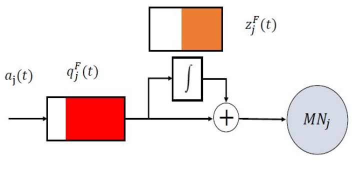
It can be observed that the **LPM** problem is naturally decoupled among DCs, and thus can be decomposed into I subproblems, each of which can be solved separately in the local DC. Each subproblem is an MILP problem with two binary variables, $(c_i^{ON}(t), d_i^{ON}(t))$. Since simultaneous charging and discharging are not allowed, there are only three combinations for the two binary variable, i.e., $(0,0), (1,0), (0,1)$. The subproblem then further boils down to three linear programming (LP) problems. The **NWR** problem can be similarly solved in a decentralized manner. Each MN solve its own **NWR**, which is an LP problem, independently with the queue information from the connected DCs. The distributed online algorithm is summarized in Algorithm 1. The large-scale stochastic MILP finally reduces to simple online LPs local to MNs and DCs.

In summary, when Algorithm 1 is executed online, the computation overhead of each DC or MN is small. Communications between the connected DC and MN are required, which is trivial compared with workload transfer. Hence, Algorithm 1 can be implemented at the cost of little computation and communication burden on MNs and DCs.

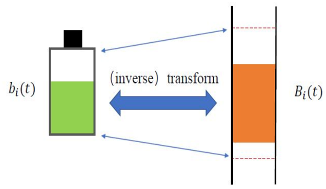
3.5 Virtual Queue Revisit

We will show in this subsection how virtual workload queues guarantee the bounded worst-case workload queuing delay and how the proper choice of parameters (V, r)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCC.2021.3132174, IEEE Transactions on Cloud Computing



(a) Virtual workload queue



(b) Virtual battery queue

Fig. 2. Relationship between the real queue and the virtual queue

ensures to respect the battery energy level constraint and maintains bounded performance gap from theoretical optimum. These are summarized in the following three theorems.

Theorem 2. *The worst-case queuing delay of the workload that arrives at MN j is bounded: delay of workload $\leq D_j^F + \max_{i\in\Omega_j^F}\{D_i^B\}$, where $D_i^B = |(Q_i^B + Z_i^B)/\varepsilon_i^B + 1|^+$ and $D_j^F = |(Q_j^F + Z_j^F)/\varepsilon_j^F + 1|^+$. $|x|^+$ denotes the smallest integer number that is larger than or equal to x. $Z_i^B = V\overline{p_i} + \varepsilon_i^B$ and $Z_j^F = V\overline{\alpha_j} + Q^{B,j} + \varepsilon_j^F$ are respectively the upper bounds of $z_i^B(t)$ and z_j^F. ($\overline{\alpha_j} \triangleq \max_{i\in\Omega_j^F}\{\alpha_{ji}\}$ and $Q^{B,j} \triangleq \max_{i\in\Omega_j^F}\{Q_i^B\}$.)*

The proof is in Appendix B (All the appendixes in the paper can be found in the supplementary materials [3]). This Theorem provides a worst-case delay guarantee for workload, in contrast with the average delay guarantee in previous studies. We briefly explain how virtual workload queues work to reduce workload delay. Assume at MN j, the workload arrival rate is low and even 0. The coefficient of $m_{ji}(t)$ in $g_{ji}(t)$, $V\alpha_{ji} + q_i^B(t) - q_j^F(t) - z_j^F(t) > 0$, is likely to hold for a long time with empty virtual workload queues, i.e., $z_i^F(t) = 0$. Since the **NWR** of MN j is to minimize $\sum_{i\in\Omega_j^F} g_{ji}(t)$, $m_{ji}(t)$ will be zero if $V\alpha_{ji} + q_i^B(t) - q_j^F(t) - z_j^F(t) > 0$, which means the workload will not be transferred for a long time, causing a large delay. However, if there is a virtual workload queue at MN j, it will keep increasing as long as there is workload waiting to be transferred until $V\alpha_{ji} + q_i^B(t) - q_j^F(t) - z_j^F(t) < 0$ for some $i \in \Omega_j^F$. Under this condition, the solution m_{ji} will be positive, and workload will be transferred. The virtual workload queues at DCs work in a similar way. As illustrated in Fig. 2(a), a virtual workload queue can be deemed as an integral of a real workload queue, which is designed to push MNs (DCs) to transfer (process) workload. Large ε_j^F and ε_i^B commit to a small queuing delay at the expense of large operational cost.

The following theorem provides a guideline to set parameters (V, r).

Theorem 3. *If $r > 1$ and V satisfies*

$$\frac{\eta^c\overline{c_i} + \frac{1}{\eta^d}\overline{d_i} + \frac{1}{r}\eta^d(M_i^B + \varepsilon_i^B)}{\eta^d(1-\frac{1}{r})\overline{p_i}} \leq V \leq \frac{\overline{b_i} - \underline{b_i}}{\eta^d\overline{p_i} - \frac{1}{r\eta^c}\underline{p_i}} \tag{27}$$

3. https://zjueducn-my.sharepoint.com/:b:/g/personal/sunjun15_zju_edu_cn/EbM6wVFi1OJLiSjID4ehNocBNWx_tZrlR2Wny8PXEaZg-Q?e=i3zDiU

the battery operation by Algorithm 1 will always respect the constraint (6).

The proof is in Appendix C. Intuitively, since the Lyapunov function is bounded all the queue lengths are finite. Meanwhile, given the proper (V, r) and limited charging and discharging rate, a virtual battery queue $B_i(t)$ only varies in a certain range, which maps exactly to $[\underline{b_i}, \overline{b_i}]$. The transformation between a real battery state and a virtual battery state is demonstrated in Fig. 2(b).

Recall that our goal is to minimize the long-term operational cost of the distributed DC network, the overall performance in terms of the cost should be evaluated. As indicated by the following theorem, the average operational cost achieved by Algorithm 1 is bounded within a constant gap from theoretical optimum.

Theorem 4. *The gap of average long-term operational cost between Algorithm 1 and theoretical optimum is bounded by:*

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t\in\mathbb{T}}\mathbb{E}[f(t)] - f^* \leq \frac{N}{V} \tag{28}$$

where $N = N_1 + rN_2$.

Proof. For brevity, let $h(\mathcal{X})$ be the right-hand side of (25) and $\mathcal{X}'$ be the optimal solution to **LPM** and **NWR**. It is straightforward that $h(\mathcal{X}') \leq h(\hat{\mathcal{X}})$ because $\mathcal{X}'$ minimizes $h(\mathcal{X})$. Therefore, the drift-plus-penalty corresponding to $\hat{\mathcal{X}}$ satisfies:

$$\Delta(t) + V\mathbb{E}[f(t)|\Theta(t)] \leq h(\mathcal{X}') \leq h(\hat{\mathcal{X}}) \leq N + V\hat{f} \tag{29}$$

where the third inequality follows from (15)-(18). Summing up (29) over $[1, 2, ..., T]$, divided by VT, and driving $T \to \infty$ result in:

$$\lim_{T\to\infty}\frac{1}{T}\sum_{t\in\mathbb{T}}\mathbb{E}[f(t)] - \hat{f} \leq \frac{N}{V} \tag{30}$$

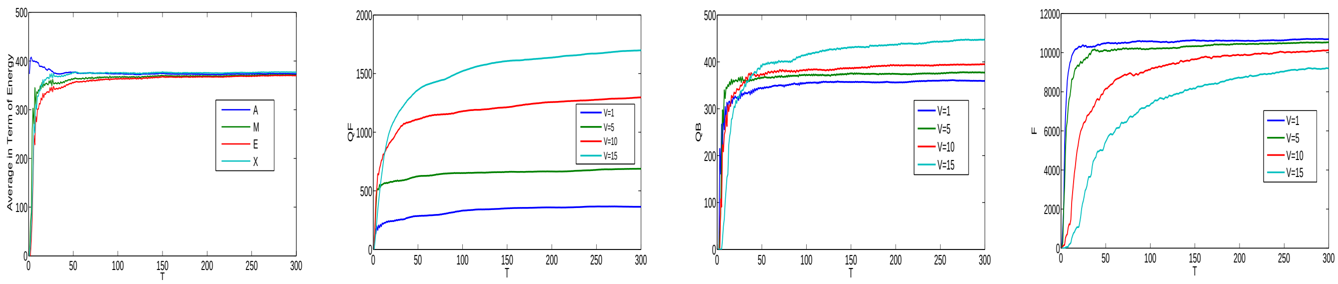
which, together with the fact $\hat{f} \leq f^*$, yields the desired gap (28). □

4 Numerical Results

Our experiments encompass two parts, synthetic simulation and simulation with realistic workload and electricity prices, corresponding to the following two subsections. In both of the two simulations, we use Matlab to emulate the data center system schematically represented in Fig. 1. The simulations are carried out on a Macbook with 2.3GHz Intel core i5, and 8G memory. Throughout the experiments, each time slot in the simulations represents one hour.

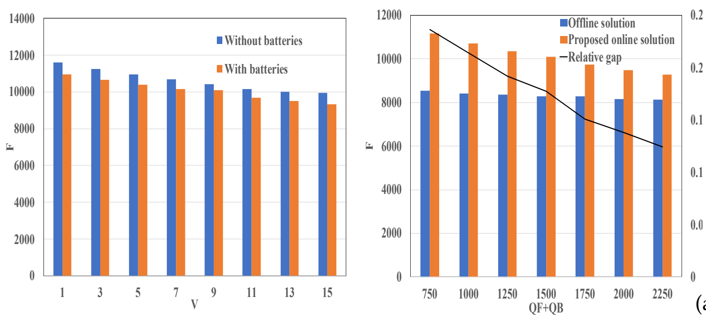
4.1 Synthetic Simulation

A small-scale network of distributed DCs is used to test the proposed method. The network consists 8 DCs and 50MNs, and the communication links are randomly picked. The electricity prices at different DCs and different time slots are uniformly distributed in the price set $\{p|p = 10 + 0.2i, i = 1, 2, \cdots, 50\}$, and the routing price α_{ij} is randomly preset from range $[2, 5]$. At each time slot, the amount of workload that enters each MN is uniformly sampled from $[0, 150]$ in terms of kWh. The parameters of the batteries are:$\overline{c} = 100kWh$, $\overline{d} = 100kWh$, $\underline{b_i} = 1000kWh$,



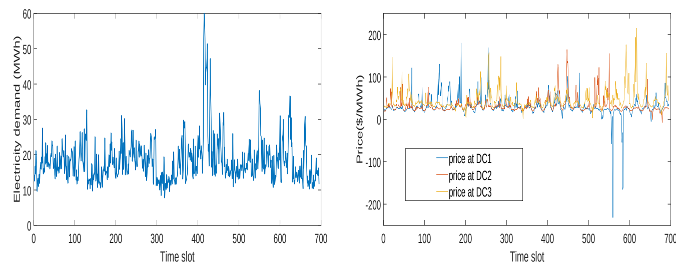
(a) Workload processing (b) The worload at MNs (c) The workload at DCs (d) The average operational cost

Fig. 3. The transition process of the system from transient state to steady state



(a) Battery's influence (b) Online and offline solution

Fig. 4. Long term average operational cost evaluation



(a) Workload trace of Google data center (b) The PJM electricity price

Fig. 5. The realistic data center workload and electricity prices

$\overline{b_i} = 2000kWh$ and $\eta^c = \eta^d = 0.8$. We set $\varepsilon_j^F = 50kWh$, $\varepsilon_i^B = 200kWh$ and $r = 2$. Some auxiliary variables are defined here to help us better describe and understand the results: Let $A = \frac{1}{T}\sum_{t\in\mathbb{T}}\sum_{j\in\mathbb{J}} a_j(t)$, $M = \frac{1}{T}\sum_{t\in\mathbb{T}}\sum_{j\in\mathbb{J}}\sum_{i\in\Omega_j^F} m_{ji}(t)$, $E = \frac{1}{T}\sum_{t\in\mathbb{T}}\sum_{i\in\mathbb{I}} e_i(t)$, and $X = \frac{1}{T}\sum_{t\in\mathbb{T}}\sum_{i\in\mathbb{I}} x_i(t)$ be the time-average total workload that arrives at all the MNs, the transferred workload, the total electricity consumption to process the workload and the electricity procured from the grid, respectively. Let $QF = \frac{1}{T}\sum_{t\in\mathbb{T}}\sum_{j\in\mathbb{J}} q_j^F(t)$, $QB = \frac{1}{T}\sum_{t\in\mathbb{T}}\sum_{i\in\mathbb{I}} q_i^B(t)$ and $F = \frac{1}{T}\sum_{t\in\mathbb{T}} f(t)$ be the time-average total workload at the MNs and DCs, and the average operational cost, respectively.

We first show the transition process of the system from transient state to steady state and the impact of V on this process and the final steady state. It is observed in Fig.3(a) that all the variables will converge to a steady-state level. It is anticipated that A, M and E will converge to the same level, and X will be above that level, due to losses in the charging and discharging process. It is seen in Fig.3(b) and Fig.3(c) that the larger V is, the more workload will be backlogged at MNs and DCs, but the lower the average operational cost will be, as shown in Fig.3(d).

Next we evaluate the benefit of batteries to cut down the operational cost and show the gap between the cost by Algorithm 1 and an offline optimum. The offline algorithm knows all the stochastic parameters in advance. It is illustrated in Fig.4(a), as V increases, the cost by Algorithm 1 decreases, and batteries reduce the cost by about 4%-6%. Fig. 4(b) suggests that both costs will decrease as the queue backlog increases, but the former decreases faster. Therefore, the relative gap defined as $RG = (F_{online} - F_{offline})/F_{offline}$ goes down from 0.230 to 0.125.

4.2 Realistic Workload and Electricity Price Simulation

For the realistic data analysis, we use the workload trace from Google cluster [45]. The workload trace represents 29 days' workload on a cluster of about 12.5k servers in May 2011. Suppose that the workload is received in 15 MNs and processed in 3 DCs, and the workload is quantified by the electricity demand, as shown in Fig 5(a), where the peak load is mapped to the full capacity of 3 DCs (60MWh). The public electricity prices in three different regions, as shown in Fig. 5(b), are drawn from the PJM ISO (About the negative electricity prices, this phenomenon indeed exists in deregulated electricity markets such as PJM, Nord pool, and CAISO. This happens when the electricity supply exceeds the demand. Because of the inertia of the generator, the generator cannot stop immediately when the supply is higher than the demand. Therefore, in order to maintain the stability of the voltage, the supply side will offer negative prices to encourage the consumer to digest extra electricity). The setup for our simulation is $\varepsilon_j^F = 0.5MWh, \forall j \in \{1, 2, \cdots, 15\}$, $\varepsilon_i^B = 2MWh$ and $r = 2$, $\overline{c_i} = 5MWh, \overline{d_i} = 5MWh, \overline{b_i} = 20MWh, \underline{b_i} = 5MWh, \eta^c = \eta^d = 0.8, E_i = 20MWh, \forall i \in \{1, 2, 3\}$, and the bandwidth limit M_{ij} and the routing cost α_{ij} between the MNs and DCs are randomly set within the ranges $[2, 5]MWh$ and $[5, 10]\$/MWh$. (As for the parameters of the data centers, they are not publicly available, and hence we set them by ourselves. For instance, $\overline{b_i}$ and $\underline{b_i}$ are determined

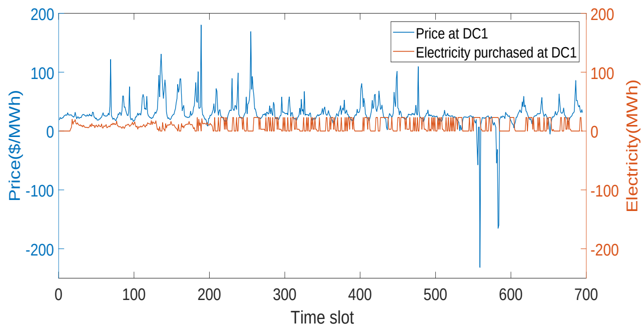


Fig. 6. The electricity procurement guided by the electricity price at DC 1

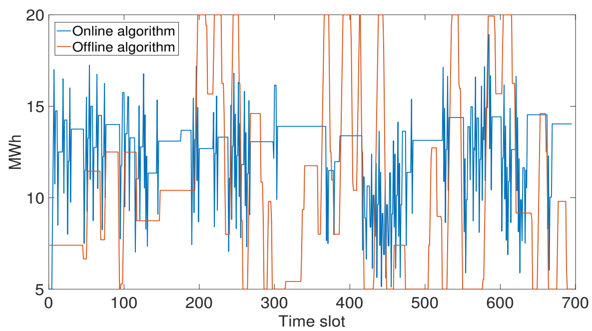


Fig. 7. The energy level of the battery at DC 1

such that when the battery is fully charged, it can support the data center for an hour [46]; while its lowest energy level can sustain the data center for about 10 to 20 minutes when power failure accidentally happens. E_i is set such that the system is able to accommodate the possible continuing high rate of incoming workload(20MWh). ϵ_j^F and ϵ_i^B are about 10% of the highest incoming workload rate at the MNs and the DCs, respectively, to penalize the accumulation of workload.)

Fig. 6 shows the electricity procurement of one of the DCs versus the local electricity prices. In general, DCs purchase electricity at low prices to process workload and charge batteries. When the price at a certain DC is high, it will not be assigned much workload to avoid purchasing a large amount of electricity. The evolution of the battery energy level at DC 1 is shown in Fig. 7. It is observed

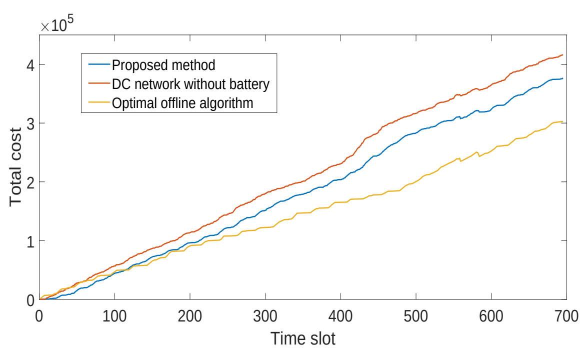


Fig. 8. The comparison of cumulative total cost

that Algorithm 1 uses the battery more frequently, and the battery constraint is always respected. Fig. 8 shows the cumulative operational cost of the DCs with and without batteries by Algorithm 1 and the offline optimal strategy. The benefit of batteries is significant by saving about 50,000 dollars (approximately 12% of the operational cost without batteries) in one month.

Fig. 6 shows the electricity procurement of one of the DCs versus the local electricity prices. The results verify that with the proposed strategy, the DCs are operating economically, meaning that they utilize cheap energy to process the workload. In specific, the DCs purchase electricity at low prices to process workload and charge batteries. When the price at a certain DC is high, this DC will not be assigned much workload to avoid purchasing a large amount of electricity. The evolution of the battery energy level at DC 1 is shown in Fig. 7, and it is shown that the energy level maintains within [-5, 20], validating Theorem 3. Moreover, it is observed that online Algorithm 1 uses the battery more frequently compared with its offline counterpart which possesses all future information, thus avoiding short-sight decisions and optimizing entire-horizon operation. Fig.8 reports the cumulative operational cost of the DCs with and without batteries by Algorithm 1 and the offline optimal strategy. The benefit of batteries is significant as it saves about 50,000 dollars (approximately 12% of the operational cost without batteries) in one month.

It is noteworthy that Algorithm 1 performs better than the offline optimal strategy in the first 100 time slots. Recall that the initial workload queue lengths are 0's. Algorithm 1 tends to postpone processing workload at the beginning, while the offline optimal strategy processes workload from a global perspective and avoids overusing the deferrability of workload at the beginning.

5 Discussions and Conclusions

5.1 Discussions

In this paper, we mainly concentrate on the operational cost minimization of distributed DCs by exploiting the deferability of workload, workload distribution and batteries. In our problem formulation some assumptions are adopted to simplify the model. These simplifications help us to ignore tedious technical details and focus our attention on our interests. We have some discussions on the assumptions below.

- The workload is quantified in terms of energy. To be specific, resource demand of workload consists of many components, including the CPU, memory, storage disk. Some models use one of the demand to represent the resource demand of the workload, and there are literatures that treat the multiple resource requests separately [47]. The electricity consumption of a data center comes from the IT infrastructure, cooling, and power conversion et al [48]. Since we concern the energy consumption of the data centers, we follow the routine to quantify the workload in terms of electricity demand [18], [49]. It should be noted that the document of Google does no quantify the workload load by energy. In the simulation, we count the tasks in each time slot and map the number of the tasks to the energy request, supposing that the energy consumption is proportional to the number of tasks [19], [49].

- The transfer cost (bandwidth cost) is proportional to the amount of transferred workload. For each communication link the transfer price is different and constant. The nonlinear convex cost is also employed in the some research, such as [11]. We employ the linear cost model because it is more brief in the theoretical analysis. It worths mentioning that the proposed method applies to the more general case where the cost is an increasing convex function.
- The stochastic workload and price follow independent and identical distribution (I.I.D.). I.I.D. is a widely used assumption in similar studies [11], [32]–[34], [50]–[55]. Therefore, we just follow the standard assumption in this field. In the simulation section, we first conduct synthetic simulation that employs workload and electricity price generated from I.I.D.. The results verify the theoretical analysis. In the real world data simulation, we use the Goolge workload trace and PJM electricity price. These real data may not rigorously follow I.I.D. The numerical results show that the proposed method performs well on these real data.

The robustness of the proposed approach can be illuminated from two perspectives. *1) The assumption:* Albeit the system is faced with stochastic workload and electricity prices, this work does not assume that the distributions of the stochastic variables are known. This makes the algorithm design and analysis distribution-free, which means the resultant algorithm is robust and applicable to a broad range of scenarios with different distributions. *2) The results:* The delay guarantee established in Theorem 2 is derived via worst-case analysis, and thus the delay guarantee is robust. As for the cost, Theorem 4 ensures that the long-term cost will not be too high relative to the optimal cost. There indeed exists such case in theory that the cost for some period might be large. But this does not undermine the value the the long-term cost guarantee, because in practice the data centers run for a long time.

5.2 Conclusions

In this paper we investigate the online operation of battery-assisted distributed DCs to minimize their long-term operational cost. We take full advantage of spatio-temporal variations of electricity prices, battery and deferrability of workload to mitigate the impact of uncertainties in future workload and electricity prices. Based on Lyapunov optimization, we design virtual queues for workload and batteries and develop a fully distributed online algorithm that is readily implementable online with provable performance guarantee. Simulations on real-world data verify our analysis and show significant cost reduction by our algorithm as well as huge savings by incorporating batteries.

References

[1] J. Struckmeier, "Data center costs as a driving force for energy efficiency: Part 1." http://www.datacenterknowledge.com/industry-perspectives/data-center-costs-driving-force-energy-efficiency-part-1. Oct 17, 2017.

[2] H. Hu, Y. Wen, L. Yin, L. Qiu, and D. Niyato, "Coordinating workload scheduling of geo-distributed data centers and electricity generation of smart grid," *IEEE Transactions on Services Computing*, 2017.

[3] M. H. Ahmed, C. Fei, F. C. Lee, and Q. Li, "48v voltage regulator module with pcb winding matrix transformer for future data centers," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 12, pp. 9302–9310, 2017.

[4] Q. Tang, S. K. S. Gupta, and G. Varsamopoulos, "Energy-efficient thermal-aware task scheduling for homogeneous high-performance computing data centers: A cyber-physical approach," *IEEE Transactions on Parallel and Distributed Systems*, vol. 19, pp. 1458–1472, Nov 2008.

[5] Q. Fang, J. Wang, Q. Gong, and M. Song, "Thermal-aware energy management of an hpc data center via two-time-scale control," *IEEE Transactions on Industrial Informatics*, vol. 13, pp. 2260–2269, Oct 2017.

[6] P. G. J. Leelipushpam and J. Sharmila, "Live vm migration techniques in cloud environment: A survey," in *2013 IEEE Conference on Information Communication Technologies*, pp. 408–413, April 2013.

[7] K. Ye, Z. Wu, C. Wang, B. B. Zhou, W. Si, X. Jiang, and A. Y. Zomaya, "Profiling-based workload consolidation and migration in virtualized data centers," *IEEE Transactions on Parallel and Distributed Systems*, vol. 26, pp. 878–890, March 2015.

[8] J. Tu, L. Lu, M. Chen, and R. K. Sitaraman, "Dynamic provisioning in next-generation data centers with on-site power production," in *International Conference on Future Energy Systems*, pp. 137–148, 2013.

[9] J. Yao, X. Liu, and C. Zhang, "Predictive electricity cost minimization through energy buffering in data centers," *IEEE Transactions on Smart Grid*, vol. 5, pp. 230–238, Jan 2014.

[10] Y. Yao, L. Huang, A. B. Sharma, L. Golubchik, and M. J. Neely, "Power cost reduction in distributed data centers: A two-time-scale approach for delay tolerant workloads," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, pp. 200–211, Jan 2014.

[11] T. Chen, A. Mokhtari, X. Wang, A. Ribeiro, and G. B. Giannakis, "Stochastic averaging for constrained optimization with application to online resource allocation," *IEEE Transactions on Signal Processing*, vol. 65, pp. 3078–3093, June 2017.

[12] D. Wang, C. Ren, A. Sivasubramaniam, B. Urgaonkar, and H. Fathy, "Energy storage in datacenters: what, where, and how much?," in *ACM SIGMETRICS Performance Evaluation Review*, vol. 40, pp. 187–198, ACM, 2012.

[13] "How big batteries at data centers could replace power plants." https://www.scientificamerican.com/article/how-big-batteries-at-data-centers-could-replace-power-plants/.

[14] W. Zheng, K. Ma, and X. Wang, "Hybrid energy storage with supercapacitor for cost-efficient data center power shaving and capping," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, no. 4, pp. 1105–1118, 2017.

[15] B. Aksanli, "Datacenter peak power management with energy storage devices," *IEEE Internet Computing*, 2017.

[16] Y. Shi, B. Xu, D. Wang, and B. Zhang, "Using battery storage for peak shaving and frequency regulation: Joint optimization for superlinear gains," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2882–2894, 2018.

[17] A. N. Toosi and R. Buyya, "A fuzzy logic-based controller for cost and energy efficient load balancing in geo-distributed data centers," in *2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing (UCC)*, pp. 186–194, Dec 2015.

[18] Y. Zhang, L. Deng, M. Chen, and P. Wang, "Joint bidding and geographical load balancing for datacenters: Is uncertainty a blessing or a curse?," *IEEE/ACM Transactions on Networking*, vol. 26, no. 3, pp. 1049–1062, 2018.

[19] P. Wang, Y. Zhang, L. Deng, M. Chen, and X. Liu, "Second chance works out better: Saving more for data center operator in open energy market," in *Information Science and Systems (CISS), 2016 Annual Conference on*, pp. 378–383, IEEE, 2016.

[20] T. Zhao and Z. Ding, "Distributed finite-time optimal resource management for microgrids based on multi-agent framework," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 8, pp. 6571–6580, 2018.

[21] A. Khosravi, A. Nadjaran Toosi, and R. Buyya, "Online virtual machine migration for renewable energy usage maximization in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCC.2021.3132174, IEEE Transactions on Cloud Computing

geographically distributed cloud data centers," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 18, p. e4125, 2017.

[22] D. Paul, W.-D. Zhong, and S. K. Bose, "Demand response in data centers through energy-efficient scheduling and simple incentivization," *IEEE Systems Journal*, vol. 11, no. 2, pp. 613–624, 2017.

[23] Q. Zhang, W. Deng, and G. Li, "Stochastic control of predictive power management for battery/supercapacitor hybrid energy storage systems of electric vehicles," *IEEE Transactions on Industrial Informatics*, 2017.

[24] M. Dabbagh, B. Hamdaoui, A. Rayes, and M. Guizani, "Shaving data center power demand peaks through energy storage and workload shifting control," *IEEE Transactions on Cloud Computing*, vol. PP, no. 99, pp. 1–1, 2017.

[25] L. J. Cupelli, T. Schutz, P. Jahangiri, M. Fuchs, A. Monti, and D. Muller, "Data center control strategy for participation in demand response programs," *IEEE Transactions on Industrial Informatics*, pp. 1–1, 2018.

[26] W. Zheng, K. Ma, and X. Wang, "Hybrid energy storage with supercapacitor for cost-efficient data center power shaving and capping," *IEEE Transactions on Parallel and Distributed Systems*, vol. 28, pp. 1105–1118, April 2017.

[27] J. Wan, B. Chen, S. Wang, M. Xia, D. Li, and C. Liu, "Fog computing for energy-aware load balancing and scheduling in smart factory," *IEEE Transactions on Industrial Informatics*, 2018.

[28] E. Gelenbe and Y. Zhang, "Performance optimization with energy packets," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3770–3780, 2019.

[29] Y. M. Kadioglu and E. Gelenbe, "Product-form solution for cascade networks with intermittent energy," *IEEE Systems Journal*, vol. 13, no. 1, pp. 918–927, 2018.

[30] E. Gelenbe and E. T. Ceran, "Energy packet networks with energy harvesting," *Ieee Access*, vol. 4, pp. 1321–1331, 2016.

[31] L. Huang and M. J. Neely, "Delay reduction via lagrange multipliers in stochastic network optimization," *IEEE Transactions on Automatic Control*, vol. 56, no. 4, pp. 842–857, 2011.

[32] M. J. Neely, "Optimal energy and delay tradeoffs for multiuser wireless downlinks," *IEEE Transactions on Information Theory*, vol. 53, pp. 3095–3113, Sept 2007.

[33] X. Wei and M. J. Neely, "Delay optimal power aware opportunistic scheduling with mutual information accumulation," *Eprint Arxiv*, vol. 57, no. 8, pp. 5151 – 5162, 2015.

[34] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends in Networking*, vol. 1, no. 1, pp. 1–144, 2006.

[35] W. Deng, F. Liu, H. Jin, C. Wu, and X. Liu, "Multigreen: cost-minimizing multi-source datacenter power supply with online control," pp. 149–159, 2013.

[36] Y. Guo, M. Pan, Y. Gong, and Y. Fang, "Dynamic multi-tenant coordination for sustainable colocation data centers," *IEEE Transactions on Cloud Computing*, pp. 1–1, 2017.

[37] E. Baccarelli, P. G. V. Naranjo, M. Shojafar, and M. Scarpiniti, "Q*: Energy and delay-efficient dynamic queue management in tcp/ip virtualized data centers," *Computer Communications*, vol. 102, pp. 89–106, 2017.

[38] M. Shojafar, C. Canali, R. Lancellotti, and E. Baccarelli, "Minimizing computing-plus-communication energy consumptions in virtualized networked data centers," in *Computers and Communication (ISCC), 2016 IEEE Symposium on*, pp. 1137–1144, IEEE, 2016.

[39] M. Alresaini, K. L. Wright, B. Krishnamachari, and M. J. Neely, "Backpressure delay enhancement for encounter-based mobile networks while sustaining throughput optimality," *IEEE/ACM Transactions on Networking*, vol. 24, pp. 1196–1208, April 2016.

[40] https://www.pjm.com/about-pjm.aspx.

[41] L. Georgiadis, M. J. Neely, and L. Tassiulas, "Resource allocation and cross-layer control in wireless networks," *Foundations and Trends® in Networking*, vol. 1, no. 1, pp. 1–144, 2006.

[42] M. J. Neely, "Energy optimal control for time-varying wireless networks," *IEEE Transactions on Information Theory*, vol. 52, pp. 2915–2934, July 2006.

[43] M. J. Neely, "Energy optimal control for time-varying wireless networks," *IEEE transactions on Information Theory*, vol. 52, no. 7, pp. 2915–2934, 2006.

[44] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.

[45] Google, "Google clusterdata2011_2traces." https://github.com/google/cluster-data/blob/master/ClusterData2011_2.md.

[46] S. R. Awasthi, S. Chalise, and R. Tonkoski, "Operation of datacenter as virtual power plant," in *Energy Conversion Congress and Exposition (ECCE), 2015 IEEE*, pp. 3422–3429, IEEE, 2015.

[47] T. H. Nguyen, M. Di Francesco, and A. Yla-Jaaski, "Virtual machine consolidation with multiple usage prediction for energy-efficient cloud data centers," *IEEE Transactions on Services Computing*, 2017.

[48] M. Dayarathna, Y. Wen, and R. Fan, "Data center energy consumption modeling: A survey," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 1, pp. 732–794, 2016.

[49] M. A. Islam, S. Ren, G. Quan, M. Z. Shakir, and A. V. Vasilakos, "Water-constrained geographic load balancing in data centers," *IEEE Transactions on Cloud Computing*, vol. 5, no. 2, pp. 208–220, 2017.

[50] Y. Yao, L. Huang, A. B. Sharma, L. Golubchik, and M. J. Neely, "Power cost reduction in distributed data centers: A two-time-scale approach for delay tolerant workloads," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 1, pp. 200–211, 2014.

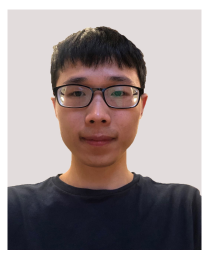
[51] M. J. Neely, "Stability and probability 1 convergence for queueing networks via lyapunov optimization," *Journal of Applied Mathematics*, vol. 2012, 2012.

[52] L. Yu, T. Jiang, and Y. Zou, "Distributed real-time energy management in data center microgrids," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3748–3762, 2018.

[53] X. Wei and M. J. Neely, "Data center server provision: Distributed asynchronous control for coupled renewal systems," *IEEE/ACM Transactions on Networking (TON)*, vol. 25, no. 4, pp. 2180–2194, 2017.

[54] T. Zhang, F. Ren, and R. Shu, "Towards stable flow scheduling in data centers," *IEEE Transactions on Parallel and Distributed Systems*, 2018.

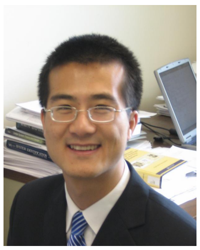
[55] W. Li, X. Zhou, K. Li, H. Qi, and D. Guo, "Trafficshaper: Shaping inter-datacenter traffic to reduce the transmission cost," *IEEE/ACM Transactions on Networking*, 2018.



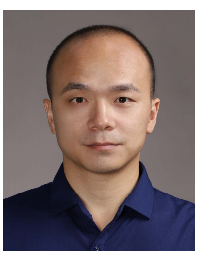
Jun Sun (S'16) received the B.S. degree in College of Astronautics from Nanjing University of Aeronautics and Astronautics, China, in 2015. Currently, he is pursuing the Ph.D. degree in the College of Control Science and Engineering, at Zhejiang University. He is a member of the Group of Networked Sensing and Control in the State Key Laboratory of Industrial Control Technology at Zhejiang University. His research interests include game theory, stochastic optimization and distributed optimization with applications in electricity market, data centers, and machine learning.

Shibo Chen received the B.Eng. degree in electronic engineering from the University of Science and Technology of China, Hefei, China, in 2011, and the Ph.D. degree in electronic and computer engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2017. He was a Postdoctoral Fellow with the Hong Kong University of Science and Technology before joining the Department of Mechanical and Energy Engineering, Southern University of Science and Technology, Shenzhen, China, in 2019 as a Research Assistant Professor. His current research interests include smart grid, optimization theory and game theory.

Pengcheng You (S'14) earned his B.S. degree (with honors) in electrical engineering and Ph.D. degree in control, both from Zhejiang University, China, in 2013 and 2018, respectively. Currently he is a Postdoctoral Fellow with the Whiting School of Engineering (ME & ECE) at Johns Hopkins University. Previously he was a visiting student at California Institute of Technology and also a research intern at Pacific Northwest National Laboratory. His research focuses on smart grid and power system economics.



Qinmin Yang received the Bachelor's degree in Electrical Engineering from Civil Aviation University of China, Tianjin, China in 2001, the Master of Science Degree in Control Science and Engineering from Institute of Automation, Chinese Academy of Sciences, Beijing, China in 2004, and the Ph.D. degree in Electrical Engineering from the University of Missouri-Rolla, MO USA, in 2007. From 2007 to 2008, he was a Post-doctoral Research Associate at University of Missouri-Rolla. From 2008 to 2009, he was a system engineer with Caterpillar Inc. From 2009 to 2010, he was a Post-doctoral Research Associate at University of Connecticut. Since 2010, he has been with the State Key Laboratory of Industrial Control Technology, the College of Control Science and Engineering, Zhejiang University, China, where he is currently a professor. He has also held visiting positions in University of Toronto and Lehigh University. He has been serving as an Associate Editor for IEEE Transactions on Systems, Man, and Cybernetics: Systems, and Automatica Sinica. His research interests include intelligent control, renewable energy systems, smart grid, and industrial big data.



Zaiyue Yang (M'10) received the B.S. and M.S. degrees from the Department of Automation, University of Science and Technology of China, Hefei, China, in 2001 and 2004, respectively, and the Ph.D. degree from the Department of Mechanical Engineering, University of Hong Kong, in 2008. He was a Postdoctoral Fellow and Research Associate with the Department of Applied Mathematics, Hong Kong Polytechnic University, before joining the College of Control Science and Engineering, Zhejiang University, Hangzhou, China, in 2010. Then, he joined the Department of Mechanical and Energy Engineering, Southern University of Science and Technology, Shenzhen, China, in 2017. He is currently a Professor there. His current research interests include smart grid, signal processing and control theory. Prof. Yang is an associate editor for the IEEE Transactions on Industrial Informatics.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCC.2021.3132174, IEEE Transactions on Cloud Computing

Supplementary materials for "Battery-Assisted Online Operation of Distributed Data Centers with Uncertain Workload and Electricity Prices"

Appendix A
The stability of the network

Before proving our theorems, we first provide the following lemma, which indicates the Lyapunov function is bounded, thus the network is stable.

Lemma 2. *All the queues in the network have a limited length, which are specified as*

$$q_i^B(t) \le V\overline{p_i} + M_i^B \triangleq Q_i^B; \tag{31a}$$
$$z_i^B(t) \le V\overline{p_i} + \varepsilon_i^B \triangleq Z_i^B; \tag{31b}$$
$$q_i^B(t) + z_i^B(t) \le V\overline{p_i} + M_i^B + \varepsilon_i^B \triangleq U_i^B; \tag{31c}$$
$$q_j^F(t) \le V\overline{\alpha_j} + \overline{Q^{B,j}} + A_j \triangleq Q_j^F; \tag{31d}$$
$$z_j^F(t) \le V\overline{\alpha_j} + \overline{Q^{B,j}} + \varepsilon_j^F \triangleq Z_i^F; \tag{31e}$$

where $\alpha_j = \max_{i\in\Omega_j^F}\{\alpha_{ji}\}$ *and* $\overline{Q^{B,j}} = \max_{i\in\Omega_j^F}\{Q_i^B\}$.

The proof is in Appendix D.

Appendix B
Proof of Theorem 2

Proof. First, we show that a request arriving at DC i will wait for at most D_i^B before it is finished:

Suppose a request arrives at t_0, and it is finished exactly at $t_0+\tau$. Since in the time duration $[t_0, t_0+\tau-1]$, the request is not finished, $q_i^B(t) > 0$ holds $\forall t \in [t_0+1, t_0+\tau]$. Therefore,

$$z_i^B(t+1) \ge z_i^B(t) - e_i(t) + \varepsilon_i^B, \forall t \in [t_0+1, t_0+\tau]$$

Summing the above inequality over $t \in [t_0+1, t_0+\tau-1]$ yields: $z_i^B(t_0+\tau) - z_i^B(t_0+1) \ge -\sum_{t=t_0+1}^{t_0+\tau-1} e_i(t) + (\tau-1)\varepsilon_i^B$. Since $0 \le z_i^B(t) \le Z_i^B$ (we omit the proof which is not difficult), we have:

$$\sum_{t=t_0+1}^{t_0+\tau-1} e_i(t) \ge (\tau-1)\varepsilon_i^B - Z_i^B \tag{32}$$

The request enters the queue at t_0+1 and is not completely removed until $t_0+\tau$. The DC processes the request in a fist-in-first-service manner, thus the electricity the DC provides during $[t_0+1, t_0+\tau-1]$ must be less than Q_i^B which is the upper bound of $q_i^B(t)$. That is,

$$\sum_{t=t_0+1}^{t_0+\tau-1} e_i(t) < Q_i^B \tag{33}$$

Combining (32) and (33) gives $\tau < \frac{Q_i^B + Z_i^B}{\varepsilon_i^B} + 1$, which means that the maximum waiting time is D_i^B.

Similarly, we can prove that the largest delay of the request at MN j is D_j^F, where $D_j^F = |(Q_j^F + Z_j^F)/\varepsilon_j^F + 1|^+$. Consequently, in the worst case, the delay of a request entering MN j is $D_j^F + \max_{i\in\Omega_j^F}\{D_i^B\}$. □

Appendix C
Proof of Theorem 3

Proof. The proof of Theorem 3 is based on the following lemma, which indicates that when the electricity of the virtual battery exceeds a certain level, the battery will not be charged any more; when it is below a certain level, the battery will not discharge any more.

Lemma 3. *The charging and discharging process of the battery follows the rule: when* $B_i(t) \ge -\frac{1}{r\eta^c}V\underline{p_i}, c_i(t) = 0$; *when* $B_i(t) \le -\frac{1}{r}\eta^d U_i^B, d_i(t) = 0$.

The proof is in Appendix E.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCC.2021.3132174, IEEE Transactions on Cloud Computing

According to the definition of $B_i(t)$, to prove the theorem, we only need to prove $B_i(t)$ satisfies the following inequality for any V satisfying (27). (Note that the number of equations in the attachment follows that in the paper. Here equation (27) is in the submitted paper.)

$$-\eta^d V\overline{p_i} + \eta^c\overline{c_i} \le B_i(t) \le \overline{b_i} - \underline{b_i} - \eta^d\overline{p_i}V + \eta^c\overline{c_i} \tag{34}$$

Obviously, $B_i(1)$ satisfies (34). Assume at time slot t, $B_i(t)$satisfies (34), then we need to prove at time slot $t+1$, $B_i(t+1)$ is also in the same range. We discuss the value of $B_i(t+1)$ given $B_i(t)$ in the following three cases:

1) If $-\frac{1}{r\eta^c}V\underline{p_i} \le B_i(t) \le \overline{b_i} - \underline{b_i} - \eta^d\overline{p_i}V + \eta^c\overline{c_i}$, then $c_i(t)=0$, thus $B_i(t+1) \le B_i(t)$. And $B_i(t+1) \ge -\frac{1}{r\eta^c}V\underline{p_i} - \frac{1}{\eta^d}\overline{d_i} \ge -\eta^d\overline{p_i}V + \eta^c\overline{c_i}$, where the second inequality holds because $V \ge \frac{\eta^c\overline{c_i}+\frac{1}{\eta^d}\overline{d_i}+\frac{1}{r}\eta^d(M_i^B+\varepsilon_i^B)}{\eta^d(1-\frac{1}{r})\overline{p_i}} \ge \frac{\eta^c\overline{c_i}+\frac{1}{\eta^d}\overline{d_i}}{\eta^d\overline{p_i}-\frac{1}{r\eta^c}\underline{p_i}}$, where the last inequality is true because of the premise that the performance of the battery satisfies (5).
2) If $-\frac{1}{r}\eta^d U_i^B \le B_i(t) < -\frac{1}{r\eta^c}V\underline{p_i}$, obviously, $B_i(t+1) \ge -\frac{1}{r}\eta^d U_i^B - \frac{1}{\eta^d}\overline{d_i} \ge -\eta^d\overline{p_i}V + \eta^c\overline{c_i}$, where the second inequality holds because of the left part inequality in (27). And $B_i(t+1) \le -\frac{1}{\eta^c}V\underline{p_i} + \eta^c\overline{c_i} \le \overline{b_i} - \underline{b_i} - \eta^d\overline{p_i}V + \eta^c\overline{c_i}$, where the second inequality holds because of the right part inequality in (27).
3) If $-\eta^d U_i^B - \frac{1}{\eta^d}\overline{d_i} \le B_i(t) < -\eta^d U_i^B$, then $d_i(t)=0$, thus $B_i(t+1) \ge B_i(t)$, and $B_i(t+1) \le -\eta^d U_i^B + \eta^c\overline{c_i} \le \overline{b_i} - \underline{b_i} - \eta^d\overline{p_i}V + \eta^c\overline{c_i}$.

Therefore, $B_i(t+1)$ is within the same area as $B_i(t)$.

From all above, it can be concluded that (34) holds for $t \in \mathbb{T}$ provided that V satisfies (27). □

Appendix D
Proof of Lemma 2

Proof. It is true that (31a) holds for $t=1$. Suppose that (31a) holds at t, then we only need to show that it also hold at $t+1$. We discuss the situation of $q_i^B(t+1)$ according to the following two cases:
(1) If $q_i^B(t) \le V\overline{p_i}$, from (2) we know that queue $q_i^B(t)$ grows up at most M_i^B in a single time slot, thus $q_i^B(t+1) \le V\overline{p_i} + M_i^B$;
(2) If $V\overline{p_i} < q_i^B(t) \le V\overline{p_i} + M_i^B$, we write the **LPM** problem of the ith DC as:

$$\begin{aligned}\textbf{LPM}_i : \min\ & g_i(t)\\ s.t.\ \ & (7)-(10),\\ & x_i(t)+d_i(t)-c_i(t) \le S_1,\ \ (\lambda_i^a)\\ & x_i(t)+d_i(t)-c_i(t) \ge 0 \quad (\lambda_i^b)\end{aligned}$$

where $S_1 = \min\{q_i^B(t) + \sum_{j\in\Omega_i^B} m_{ji}(t), E_i\}$ and (7)-(10) is only for t and i. Let λ_i^a and λ_i^b be the Lagrangian multipliers associated with the corresponding constraints. The primal-dual optimizer $(x_i(t), \lambda_i^a, \lambda_i^a)$ complies with the following KKT conditions:

$$Vp_i - q_i^B(t) - z_i^B(t) + \lambda_i^a - \lambda_i^b = 0; \tag{35a}$$
$$\lambda_i^a \ge 0, \lambda_i^b \ge 0; \tag{35b}$$
$$(x_i(t)+d_i(t)-c_i(t)-S_1)\lambda_i^a = 0; \tag{35c}$$
$$(x_i(t)+d_i(t)-c_i(t)-S_1)\lambda_i^b = 0. \tag{35d}$$

It is clear that $\lambda_i^a = q_i^B(t) + z_i^B(t) - Vp_i + \lambda_i^b > 0$. Therefore, it can be derived from (35c) and (11) that $e_i(t) = x_i(t) + d_i(t) - c_i(t) = S_1$. If $e_i(t) = S_1 = q_i^B(t) + \sum_{j\in\Omega_i^B} m_{ji}(t)$, then $q_i^B(t+1)=0$; If $e_i(t) = S_1 = E_i$, then $q_i^B(t+1) \le q_i^B(t) \le V\overline{p_i} + M_i^B$.

Here we complete the proof of (31a), and similarly (31b) and (31c) can be proved. The proof of (31d) and (31e) is roughly the same as the above, following the mathematical induction fashion. Therefore, we omit the main body of the proof and only prove that when

$$q_j^F(t) > V\overline{\alpha_j} + \overline{Q^{B,j}}, \tag{36}$$

$\sum_{i\in\Omega_j^F} m_{ji}(t) = \min\{q_j^F(t) + a_j(t), M_j^F\} \triangleq S_2$ must hold, which indicates that when the workload backlogged at the MN exceeds a threshold, the MN will process the workload at a largest possible rate.

The network workload routing problem for MN j is

$$\begin{aligned}\textbf{NWR}_j : \min & \sum_{i\in\Omega_j^F} g_{ij}(t)\\ s.t.\quad & -m_{ji}(t) \le 0, \forall i \in \Omega_i^F \ \ (\lambda_i^l)\\ & m_{ji}(t) \le M_{ji}, \forall i \in \Omega_j^F \ \ (\lambda_i^r)\\ & \sum_{i\in\Omega_j^F} m_{ji}(t) \le S_2 \qquad (\lambda)\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCC.2021.3132174, IEEE Transactions on Cloud Computing

Let λ_i^l, λ_i^r and λ be the Lagrangian multipliers associated with the three constraints. The primal-dual optimizer $(m_{ji}(t), \lambda_i^l, \lambda_i^r, \lambda)$ complies with the following KKT conditions:

$$V\alpha_{ji} + q_i^B(t) - q_j^F(t) - z_j^F(t) - \lambda_i^l + \lambda_i^r + \lambda = 0; \tag{37a}$$

$$m_{ji}(t) \geq 0, m_{ji}(t) \leq M_{ji}, \sum_{i \in \Omega_j^F} m_{ji}(t) \leq S_2; \tag{37b}$$

$$\lambda_i^l \geq 0, \lambda_i^r \geq 0, \lambda \geq 0; \tag{37c}$$

$$\lambda_i^l m_{ji}(t) = 0, \lambda_i^r (m_{ji}(t) - M_{ji}) = 0; \tag{37d}$$

$$\lambda(\sum_{i \in \Omega_j^F} m_{ji}(t) - S_2) = 0. \tag{37e}$$

We discuss these conditions in two cases:
(1) $\lambda > 0$: because of (37e), $\sum_{i \in \Omega_j^F} m_{ji}(t) = S_2$ must hold.
(2) $\lambda = 0$: (37a), (37c) and (36) lead to $\lambda_i^r > 0$, which together with (37d) results in $m_{ji}(t) = M_{ji}, \forall i \in \Omega_j^F$. Therefore, $\sum_{i \in \Omega_j^F} m_{ji}(t) = M_j^F$ holds considering the fact that $M_j^F \leq \sum_{i \in \Omega_j^F} M_{ji}$. □

Appendix E
Proof of Lemma 3

Proof. Assume $B_i(t) \geq -\frac{1}{r\eta^c} V \underline{p_i}$ and $c_i(t) > 0$. Suppose the best possible amount of power drawn from the grid is $x_i(t)$ given $c_i(t) > 0$, then the corresponding cost $g_i(t)$ satisfies:

$$\begin{aligned} g_i(t) \geq & V p_i(t) x_i(t) - V \underline{p_i} c_i(t) \\ & + [q_i^B(t) + z_i^B(t)][-x_i(t) + c_i(t)] + V\beta c_i^{ON}(t) \\ \geq & [V p_i(t) - q_i^B(t) - z_i^B(t)][x_i(t) - c_i(t)] + V\beta c_i^{ON}(t) \end{aligned}$$

where the right hand side of the second inequality is exactly the cost if $x_i(t) - c_i(t)$ electricity is bought from the grid. Therefore, when $B_i(t) \geq -\frac{1}{r\eta^c} V \underline{p_i}$, $c_i(t) = 0$ must hold.

Assume $B_i(t) \leq -\frac{1}{r}\eta^d U_i^B$ and $d_i(t) > 0$. Suppose the best possible amount of power drawn from the grid is $x_i(t)$ given $d_i(t) > 0$, then the corresponding cost $g_i(t)$ satisfies:

$$\begin{aligned} g_i(t) \geq & [V p_i(t) - q_i^B(t) - z_i^B(t)] x_i(t) \\ & + [Q Z_i^B - q_i^B(t) - z_i^B(t)] d_i(t) \\ \geq & [V p_i(t) - q_i^B(t) - z_i^B(t)] x_i(t) \end{aligned}$$

where the right hand side of the second inequality is exactly the cost if $x_i(t)$ electricity is bought from the grid and no electricity is discharged. Therefore, when $B_i(t) \leq -\frac{1}{r}\eta^d U_i^B$, $d_i(t) = 0$ must hold. □