

碳感知的绿色云数据中心能源优化在线算法

何怀文^{1,2}, 肖涛¹, 程东¹, 彭政¹, 傅瑜¹

(1. 电子科技大学中山学院计算机学院 广东 中山 528402; 2. 中山大学数据科学与计算机学院 广州 510275)

【摘要】使用绿色能源的云数据中心需要满足服务水平协议(service layer agreement, SLA)和长期平均碳排放约束下的能源成本最小化问题, 针对该问题, 本文进行了深入研究和分析。通过将问题形式化为受约束的随机优化问题, 利用Lyapunov优化理论, 提出一个不需要预测未来数据的碳感知能源成本优化在线算法, 并进一步证明该在线算法可以获得接近离线最优算法的次优解。基于真实数据的仿真实验结果表明: 该算法比基准算法具有更好的性能, 并且能通过调整控制参数接近离线的最优解。

关键词 碳排放; 云数据中心; 可再生能源; 随机网络优化

中图分类号 TP393.02 **文献标志码** A **doi**:10.3969/j.issn.1001-0548.2018.04.013

Carbon-Aware Power Optimal Online Algorithm for Green Cloud Data Center

HE Huai-wen^{1,2}, XIAO Tao¹, CHENG Dong¹, PENG Zheng¹, and FU Yu¹

(1. School of Computer, University of Electronic Science and Technology of China Zhongshan Institute Zhongshan Guangdong 528402;

2. School of Data and Computer Science, Sun Yat-sen University Guangzhou 510275)

Abstract A deep insight into the problem of energy cost optimization for cloud data center using green energy under constraints of satisfying SLA and long-term average carbon emissions. By formulating the problem into a constrained stochastic optimization problem, a carbon-aware energy cost optimal online algorithm based on Lyapunov optimization theory is proposed which can work without any future information. Furthermore, it is proved that this online algorithm can achieve the close-to-minimum solution compared to the offline optimal algorithm. Results of simulation based on real world data trace show that this algorithm has better performance than benchmark algorithm and can achieve offline optimal solution with tuning control parameters.

Key words carbon emission; cloud data center; renewable energy; stochastic network optimization

近年来, 云数据中心每年所需的巨额能源费用以及产生的温室气体排放, 已引起了学术界和工业界重要关注^[1-2]。据统计, 信息和通信技术(information communication technology, ICT)行业的碳排放已占到全球碳排放的2%, 与航天工业相当, 并且预计将到2020还要翻倍^[3]。而随着可再生能源技术的发展, 以及其污染少、用之不竭等特点, 已被多个著名IT企业(如Google^[4]、Apple^[5]、Microsoft^[6])用于数据中心电力供应, 以降低二氧化碳排放。因此, 新能源的利用对构建绿色数据中心及其可持续发展, 具有重要意义。

虽然可再生能源比传统电网能源具有更低的碳排放, 但是可再生能源(尤其是风能和太阳能)受到当

地气候条件影响较大, 其能源供应具有间歇性和难以预测, 它们的引入给云数据中心能源管理和负载调度带来了更大挑战^[7]。同时, 由于可再生能源的不稳定性, 使得其电力成本往往高于传统电网电力价格。面对负载的随机性、电力价格的波动性和绿色能源的间歇性, 需要设计在线的资源调度算法, 在多种能源供应环境下以保证SLA和碳排放的约束条件, 使得数据中心成本最小化是一个颇具挑战性的问题。

为了解决以上问题, 本文从不同类型负载的特征出发, 首先将数据中心能源成本优化问题形式化为一个受约束的随机优化问题, 然后利用Lyapunov优化理论, 设计了一个碳感知的能源优化在线控制

收稿日期: 2017-07-21; 修回日期: 2017-11-15

基金项目: 国家自然科学基金(61502088); 广东省自然科学基金(2016A030313018); 广东省高等学校优秀青年教师基金(YQ2015241); 广东省青年创新项目(2015KQNCX206)

作者简介: 何怀文(1980-), 男, 博士生, 副教授, 主要从事云计算、资源调度、随机优化方面的研究。

算法, 在线决策不同能源购买量以及数据中心活动的服务器数量。最后, 本文基于真实数据进行了大量的仿真实验, 验证了算法性能。

1 相关工作

利用绿色能源和电力价格的时变性来进行资源调度, 解决数据中心高能耗、高费用、高污染等问题, 已经成为了目前的研究热点^[2, 8-9]。文献[10-11]利用电力价格的时空化特性, 将数据中心能耗优化问题形式化为混合非线性整数规划问题并进行求解。文献[12-13]扩展了文献[10-11]的模型, 提出可再生能源下的数据中心负载均衡在线算法。但上述研究仅仅考虑了响应式负载的SLA和数据中心成本的关系, 并没有考虑到数据中心碳排放的环境影响。

文献[14]关注了延迟容忍负载的优化调度问题, 基于Lyapunov优化提出一种双尺度的跨地域负载调度算法。文献[15-16]针对批处理作业的GreenHadoop和GreenSlot调度系统, 通过预测太阳能的可用量来达到最大化太阳能利用率的目的, 但以上系统并不适用于延迟敏感的应用。文献[17]通过时隙数区分不同负载类型, 设计了可以同时调度不同负载的在线能源算法。文献[18-20]通过预测未来一段时期内可再生能源的发电量, 然后延迟批处理负载的执行以降低数据中心的能源成本, 但是对于难以预测的可再生能源而言, 算法很难获得相应的性能保证。

与原有工作相比, 本文综合考虑了多种能源供应下, 数据中心性能、成本和碳排放三者之间的相互约束和平衡, 提出了一种不需要获知未来数据的最小化数据中心能源成本的在线调度算法。

2 系统模型与优化目标

本文假设云数据中心能源供应包括: 褐色能源、太阳能和风能。褐色能源以传统电网供电为主, 太阳能和风能则从当地可再生能源发电公司购买(如Google向Iowa公司购买了20年风力电能^[21])。数据中心系统模型如图1所示。

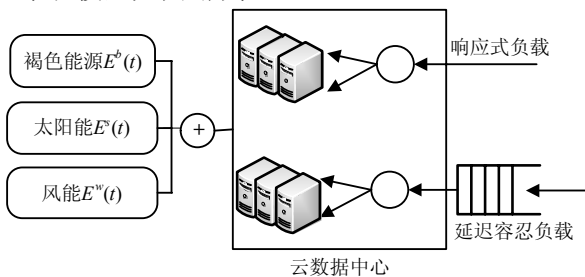


图1 绿色数据中心系统模型

2.1 负载模型

数据中心请求通常分为响应式请求和延迟容忍请求^[22]。响应式请求指实时要求较高、需要在很短时间内处理完并返回结果的用户请求, 如Web页面访问、网络游戏等, 该类负载往往占到数据中心流量的50%以上^[13, 23]。延迟容忍请求则通常为批处理任务, 需要较长处理时间, 如MapReduce^[24]和大规模图形处理^[25]等。

与文献[26]相似, 本文采用离散时隙模型, 假设在每个在时隙 $t = (1, 2, \dots, \tau, \dots)$, 到达数据中心响应式、延迟容忍负载分别记为 $W_1(t)$ 和 $W_2(t)$, 并且 $W_1(t)$ 、 $W_2(t)$ 满足约束:

$$0 \leq W_1(t) \leq W_1^{\max}, \quad 0 \leq W_2(t) \leq W_2^{\max} \quad (1)$$

式中, W_1^{\max} 、 W_2^{\max} 分别为数据中心能处理的响应式和延迟容忍负载的最大容量。

如图1所示, 处理响应式和延迟容忍负载的活动服务器数量分别为记为 $n_1(t)$ 和 $n_2(t)$, 数据中心服务器总数量记为 M , 则有:

$$n_1(t) \geq 0, \quad n_1(t) \in \mathcal{N} \quad (2)$$

$$n_2(t) \geq 0, \quad n_2(t) \in \mathcal{N} \quad (3)$$

$$n_1(t) + n_2(t) \leq M \quad (4)$$

对于响应式负载, 通常采用平均响应时间作为SLA^[12]。本文采用M/GI/1/PS排队模型来模拟响应式负载服务器群 $n_1(t)$, 假设负载被平均分发到每台处理速率为 μ_1 的服务器进行处理。根据文献[27], 得到平均响应时间为 $\frac{1}{\mu_1 - W_1(t)n_1^{-1}(t)}$, 则响应式负载的SLA约束可以表示为:

$$\frac{1}{\mu_1 - W_1(t)n_1^{-1}(t)} \leq d_{\max} \quad (5)$$

式中, d_{\max} 为云中心需要保证的最大平均响应时间。

另外, 服务器利用率 $\rho = \frac{W_1(t)}{n_1(t)\mu_1} < 1$, 以保证服务器的利用率不会过载。

延迟容忍负载允许推迟若干时隙才开始执行, 因此可以将延迟容忍负载推迟到电力价格较低时处理, 以节省能源成本。本文假设延迟容忍负载的动态队列为 $U(t)$, 该类负载服务器处理速率为 μ_2 , 则 $U(t)$ 的队长变化为:

$$U(t+1) = \max[U(t) - n_2(t)\mu_2, 0] + W_2(t) \quad (6)$$

对于队列 $U(t)$, 需要保证该队列的长期稳定性, 即保证 $U(t)$ 时间平均队长期望值小于无穷大, 有:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E[U(t)] \leq \infty \quad (7)$$

2.2 能源模型

本文只考虑风能和太阳能两种最常用的可再生能源。风能和太阳能的发电量受当时地域气候(风速和太阳光照辐射)影响较大。假设在时隙 t 风能和太阳能的发电量分别为 $E_{\text{supply}}^w(t)$ 和 $E_{\text{supply}}^s(t)$ 。根据文献[28-30]的能源模型,可得风能和太阳能的发电量表示如下:

1) 风能发电量: $E_{\text{supply}}^w(t) = \frac{1}{2} \beta B \rho v^3(t) l$ 。其中 $v(t)$ 为时隙 t 的风速, B 为所有风力涡轮机转子总面积, β 为风能到电能的转换效率系数, ρ 为空气密度, l 是时间长度。

2) 太阳能发电量: $E_{\text{supply}}^s(t) = \alpha A s(t) l$ 。其中 $s(t)$ 为时隙 t 的光照辐射强度, A 为太阳能面板总面积, α 为太阳能到电能的转换效率系数, l 是时间长度。

本文假设数据中心中每台活动服务器能耗模型相同^[12, 26],单位时间能耗为 P_0 ,则时隙 t 数据中心总能耗为 $(n_1(t) + n_2(t))P_0 \times \text{PUE}$,其中电能使用效率(power usage effectiveness, PUE)为数据中心消耗的总能源与IT设备消耗的能源之比。假设在时隙 t 数据中心购买的褐色能源、太阳能和风能分别为 $E^b(t)$ 、 $E^s(t)$ 和 $E^w(t)$,则有:

$$E^b(t) + E^s(t) + E^w(t) = (n_1(t) + n_2(t))P_0 \times \text{PUE} \quad (8)$$

另外,风能和太阳能的购买量不能超过当时的发电量,因此有:

$$0 \leq E^s(t) \leq E_{\text{supply}}^s(t) \quad (9)$$

$$0 \leq E^w(t) \leq E_{\text{supply}}^w(t) \quad (10)$$

2.3 碳排放模型

在消耗等量能源的情况下,可再生能源碳排放小于褐色能源碳排放。碳排放速率(carbon emission rate, CER)用于表示不同能源的碳排放速率。假设褐色能源、风能、太阳的CER分别记为 e_b, e_w, e_s ,具体数值如表1^[31]所示。

表1 不同能源的碳排放速率

能源类型	煤炭能源/g·KWh ⁻¹	风能/g·KWh ⁻¹	太阳能/g·KWh ⁻¹
CER	968	22.5	53

因此,时隙 t 数据中心总碳排放量表示为:

$$E(t) = e_b E^b(t) + e_s E^s(t) + e_w E^w(t) \quad (11)$$

现实中,数据中心需要控制某个时期内的碳排放量。类似文献[31],本文假设数据中心的时间平均碳排放量满足以下约束:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E[E(t)] \leq H \quad (12)$$

式中, H 为数据中心的时间平均碳排放最大限制。

2.4 问题描述

本文假设数据中心参与了电力批发市场,电力价格随着时间动态波动^[1]。时隙 t 的褐色能源、风能、太阳能电力价格分别为 $P^b(t)$ 、 $P^w(t)$ 和 $P^s(t)$ 。因此,时隙 t 数据中心能源成本可表示为:

$$C(t) = P^b(t)E^b(t) + P^s(t)E^s(t) + P^w(t)E^w(t) \quad (13)$$

记 $\vec{E}(t) = \{E^b(t), E^s(t), E^w(t)\}$, $\vec{n}(t) = \{n_1(t), n_2(t)\}$,则本文优化问题描述为:在时变的电力价格和可再生能源供应下,数据中心在每个时隙选择能源购买量 $\vec{E}(t)$ 和活动服务器数量 $\vec{n}(t)$,在满足不同负载SLA以及长期碳排放约束下,使得长期能源成本最小。该问题为一个随机优化问题,数学表示为:

$$(P1) \text{ minimize } \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E[C(t)] \quad (14)$$

$$\text{满足约束: 式(1)~式(12)} \quad (15)$$

3 问题求解

3.1 离线问题求解

在离线情况下,如获知将来每个时隙的系统状态数据(如负载、电力价格、风速和光照辐射等)后,问题P1属于一个混合整数规划(mixed-integer linear programming, MILP)问题。而由于问题规模过大,求解P1时易碰到“维度灾难”导致算法收敛慢等问题。与文献[22, 32]类似,本文设计了一个向前看 K -步的离线求解算法。假设算法运行时间段共 J 个时隙。将整个周期 J 分为 R 个时间帧,每个时间帧包含有 K 个时隙,则有 $J = RK$ 。假设可以提前获知 K 个时隙内的所有系统信息,因此,优化时间段 J 的平均成本可以转换为求解 R 个P2优化问题,其中P2描述为:

$$(P2) \text{ minimize } \frac{1}{K} \sum_{t=rK}^{(r+1)K-1} C(t) \quad (16)$$

$$\text{满足: 约束式(1)~式(11)} \quad (17)$$

$$\frac{1}{K} \sum_{t=rK}^{(r+1)K-1} E(t) \leq H \quad (18)$$

为了保证问题P2存在可行解,本文做了以下假设: 1) 边界假设: 对于 $t \in \{0, 1, \dots, J\}$,假设负载到达速率 $W_1(t)$ 、 $W_2(t)$ 是有限的; 2) 可行性假设。假设对于每一个帧 $r \in \{0, 1, \dots, R\}$,都存在至少一个可行策略,满足问题P2的所有约束。而问题P2可以松弛为一个线性规划问题,可通过相关算法求解。因此,通过求解问题P2,可以得到第 r 帧的最小平均值

为 $G_r^*, r \in \{0, 1, \dots, R\}$ 。因此, 在 J 时期内, 向前看 K -步算法得到的最小成本值为 $\frac{1}{R} \sum_{r=0}^R G_r^*$ 。

3.2 在线算法设计

在现实生活中, 信息往往难以获得或准确预测。负载、电力价格以及绿色能源供应量的随机性和动态性给问题的求解带来很大挑战。因此, 本文利用最新发展的Lyapunov优化理论^[32], 设计了一种在无预测未来数据, 仅依靠当前系统状态而做出决策的在线算法。该算法复杂性低、易于部署, 可获得接近最优解的次优解。

首先将长期碳排放约束式(12)转换为队列稳定性问题。构造“碳排放虚拟队列” $Q(t)$, 表示数据中心碳排放量与长期碳排放约束 H 的偏移, 并初始化 $Q(t)=0$ 。 $Q(t)$ 队长动态变化如下:

$$Q(t+1) = \max[Q(t) - H + E(t), 0] \quad (19)$$

如果队列 $Q(t)$ 稳定, 即 $\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E[Q(t)] \leq \infty$, 则意味式(12)可以满足^[31]。定义 $\Theta(t) \triangleq [Q(t), U(t)]$, 根据Lyapunov优化框架, 定义Lyapunov优化函数和一个时隙的Lyapunov偏移量分别为 $L(t) \triangleq \frac{1}{2}[Q^2(t) + U^2(t)]$ 和 $\Delta(t) \triangleq E[L(t+1) - L(t) | \Theta(t)]$ ^[32]。通过控制 $\Delta(t)$ 来调整队列 $Q(t), U(t)$ 的拥塞程度。另外, 将优化目标的期望值(惩罚函数)与 $\Delta(t)$ 相加, 可以得到时隙 t 的Lyapunov偏移加惩罚之和为:

$$\Delta_V(t) \triangleq \Delta(t) + VE[C(t) | \Theta(t)] \quad (20)$$

式中, $V > 0$, 为控制参数, 用于调节目标函数和约束条件之间的权重。

引理 1 对于任何可行的控制策略 $\bar{E}(t)$ 和 $\bar{n}(t)$, 一个时隙的Lyapunov偏移值 $\Delta(t)$ 的上界表示为:

$$\Delta(t) \leq B + E[Q(t)(E(t) - H) + U(t) \times (W_2(t) - n_2(t)\mu_2) | \Theta(t)] \quad (21)$$

证明: 根据不等式:

$$\max[a - b + c, 0]^2 \leq a^2 + b^2 + c^2 + 2a(c - b)$$

$\forall a, b, c \geq 0$, 可以得到:

$$\frac{Q^2(t+1) - Q^2(t)}{2} \leq \frac{H^2 + E^2(t)}{2} + Q(t)(E(t) - H) \quad (22)$$

同样根据不等式:

$$(\max[a - b, 0] + c)^2 \leq a^2 + b^2 + c^2 + 2a(c - b)$$

得到:

$$\frac{U^2(t+1) - U^2(t)}{2} \leq \frac{(n_2(t)\mu_2)^2 + (W_2(t))^2}{2} + U(t)(W_2(t) - n_2(t)\mu_2) \quad (23)$$

将式(22)和式(23)相加后, 两边取期望, 则得到引理1, 其中 $B = \frac{(M\mu_2)^2 + (W_2^{\max})^2}{2} + \frac{H^2 + E_{\max}^2}{2}$, E_{\max}

表示在某个时隙时数据中心的最大碳排放量。证毕。

根据引理1和式(20), 可得到时隙 t 的Lyapunov偏移加惩罚之和 $\Delta_V(t)$ 的上界为:

$$\Delta_V(t) \leq B + E[-Q(t)H + U(t)W_2(t) | \Theta(t)] + E[VC(t) + Q(t)E(t) - n_2(t)\mu_2 | \Theta(t)] \quad (24)$$

根据Lyapunov优化理论, 最小化每个时隙 t 的偏移加惩罚之和 $\Delta_V(t)$, 则可获得目标函数的最优解^[32]。而在式(24)中, 队长 $Q(t)$ 、 H 、 $U(t)$ 和 $W_2(t)$ 在每个时隙 t 开始可以获得, 因此, 问题P1可以转换为最小化每个时隙 t 的上界, 即转换为问题P3, 表示为:

$$(P3) \underset{\bar{E}(t), \bar{n}(t)}{\text{minimize}} VC(t) + Q(t)E(t) - n_2(t)\mu_2 \quad (25)$$

满足约束: 式(2)~式(5), 式(8)~式(11) (26)

因此, 本文提出碳感知成本最小化在线算法(carbon-aware for cost minimization online algorithm, CACM)的具体描述如下:

算法1: 碳感知成本最小化在线算法—CACM算法

输入: 负载 $W_1(t), W_1(t)$; 气候数据 $s(t), v(t)$; 电力价格 $P^b(t), P^w(t), P^s(t)$

输出: 决策变量 $\bar{E}(t)$ 和 $\bar{n}(t)$

1) 在每个时隙 t 开始时刻, 获取当前队列 $U(t)$ 和 $Q(t)$ 队长以及其他系统状态信息(负载、电力价格和气候数据)。

2) 求解优化问题P3, 满足约束条件式(2)~式(5), 式(8)~式(11), 式(13)。

3) 根据式(6)、式(19)分别更新队列 $U(t)$ 和 $Q(t)$ 的队长。

4) 更新时隙 $t: t \leftarrow t + 1$

P3目标函数中 $C(t)$ 是一个包含了 $E^b(t), E^w(t), E^s(t)$ 的线性表达式, 同时 $n_1(t)$ 可以由式(5)求得, 因此, 求解P3即可获得每个时隙 t 的决策变量 $\bar{E}(t)$ 和 $\bar{n}(t)$ 。而问题P3虽然是一个MILP问题, 但是数据中心中整数变量 $n_1(t), n_2(t)$ 通常较大, 可以将其松弛到实数域进行求解, 即问题P3可转换为一个仅具有5个变量的线性规划问题, 使用内点法在多项式时间内获得有效解, 最坏情况下时间复杂度为 $O(n^{3.5}L^2)$, 其中 $n=5, L$ 为输入长度。由此可见, CACM算法能够在线运行。本文使用了IBM CPLEX^[33]工具对问题P3进行求解。

3.3 在线算法性能分析

定理 1 (CACM算法性能) 假设系统状态

$(W_1(t), W_2(t), s(t), v(t), P^b(t), P^w(t), P^s(t))$ 在时隙内属于独立同分布, 则对于控制参数 $V > 0$, CACM算法可以达到以下性能保证:

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E[C(t)] \leq C^* + \frac{B}{V} \quad (27)$$

式中, C^* 为时间平均能源成本的理论下界。

证明: 根据Lyapunov优化理论^[32], 对于问题P1, 至少会存在一个随机平稳控制策略 π (选择 $\bar{E}(t)$ 和 $\bar{n}(t)$), 满足P1的所有约束并得到以下结果:

$$E[C(t)] = C^* \quad (28)$$

$$E[W_2(t)] \leq E[n_2 \mu_2] \quad (29)$$

$$E[E(t)] \leq H \quad (30)$$

根据式(20)和式(21), 有:

$$\Delta V(t) \leq B + E[Q(t)(E(t) - H) + U(t)(W_2(t) - n_2(t)\mu_2)\Theta(t)] + VE[C(t)\Theta(t)] \leq B + VC^*$$

根据期望迭代法则, 得到:

$$E[L(t+1) - L(t)] + VE[C(t)|\Theta(t)] < B + VC^* \quad (31)$$

将 $t = 0, 2, \dots, T-1$ 时隙相加求和, 并且两边除以 T , 得到:

$$\frac{1}{T} V \sum_{t=0}^{T-1} E[C(t)] \leq B + VC^* + \frac{E[L(0)] - E[L(T)]}{T}$$

因为 $E[L(T)]$ 有限, $E[L(T)]$ 非负, 取 $T \rightarrow \infty$, 可以得到式(27), 定理1得证。

4 实验仿真

4.1 参数设置

本文基于真实负载、气候、电力价格等数据, 模拟一个常见的互联网云数据中心。仿真实验中每个时隙 t 长度为10分钟, 共模拟为4 464个时隙(31天), 相关参数设置如下:

1) 绿色能源数据。从美国国家可再生能源实验室(national renewable energy laboratory, NREL)^[33]获取位于亚利桑那大学测量站从2015年10月1日~2015年10月31日测得的风速 $v(t)$ 和太阳光照辐射强度 $s(t)$ 的历史数据。对于绿色能源电力供应模型, 本文采用与文献[34-35]相同的参数设置, 假设太阳能和风能的电力转换效率系数分别为 $\alpha=20\%$, $\beta=30\%$ 。太阳能发电面板总面积 $A=1\ 000\ \text{m}^2$, 风力涡轮机转子总面积 $B=650\ \text{m}^2$ 。空气密度为 $\rho=1.225\ \text{kg/m}^3$ 。根据前面描述的风力和太阳能发电模型, 可得数据中心绿色能源电力供应数据。开始一周风能和太阳能发电量如图2a所示。

2) 动态电力价格数据。从纽约独立系统调度机

构(new york independent system operator, NYISO)^[36]获取该地区相应时间的实时分区边际电价(locational marginal price, LMP)作为褐色能源电力价格 $P^b(t)$, 并将实时电力价格缩放到1个时隙时间。类似文献[21], 本文假设风能和太阳能电力价格分别为 $P^b(t)+1.5$ 美分和 $P^b(t)+18$ 美分。开始一周风能和太阳能电力价格变化如图2b所示。

3) 负载数据。本文同样选取从2015年10月1日开始的Wiki Dump^[37]访问量作为数据中心负载样本。延迟容忍负载占总负载比例为随机数 γ , 假设 γ 服从 $0.2 \sim 0.4$ 的均匀分布, 即 $\gamma \sim U(0.2, 0.4)$ ^[22]。开始一周负载流量如图2c所示。

4) 数据中心参数设置。假设该数据中心拥有服务器总量 $M=2\ 000$, 服务器功率为230 W, 数据中心PUE=1.2。假设响应式负载平均响应时间限制 $d_{\max}=0.1\ \text{s}$, 服务速率 $\mu_1=20$, 延迟容忍负载服务速率 $\mu_2=0.05$ 。

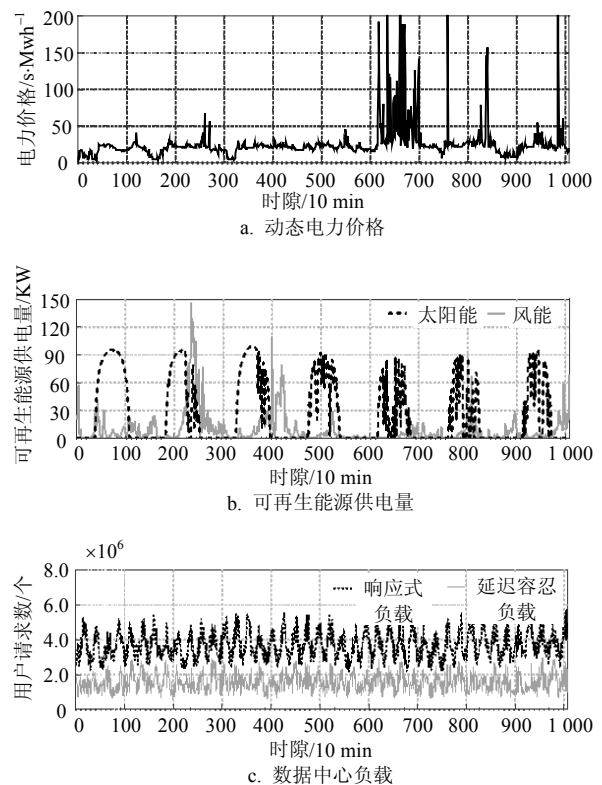


图2 实验基本参数

4.2 基准参数

本文使用以下算法作为数据中心能源成本和碳排放的基准算法: 1) 最小化数据中心能源成本Min-Cost算法; 2) 最小化碳排放Max-Green算法。其中Min-Cost算法假设只使用价格最低的褐色能源, 没有采用绿色能源; Max-Green则优先采用碳排

放率低的绿色能源,在能源不足时才使用褐色能源。Min-Cost算法和Max-Green算法获得数据中心平均能源成本和平均碳排放边界值如表2所示。

表2 数据中心平均能源成本和平均碳排放边界值

数值	方法	
	Max-Green	Min-Cost
平均碳排放/kg	46.73	89.05
平均成本/\$	6.89	2.38

4.3 结果讨论

1) 控制参数V的影响

假设长期碳排放限制为 $H=60\text{ kg}$ 。数据中心平均能源成本值 $C(t)$ 随着时隙 t 的变化曲线如图3所示。由图3可见,平均能源成本 $C(t)$ 在开始阶段处于不稳定状态,而随着时隙增加, $C(t)$ 逐渐变得平稳,说明本文算法可使平均成本逐渐趋于平稳。另外,在图3中,当控制参数 V 增大时,数据中心平均成本将越小,与定理1相符合。

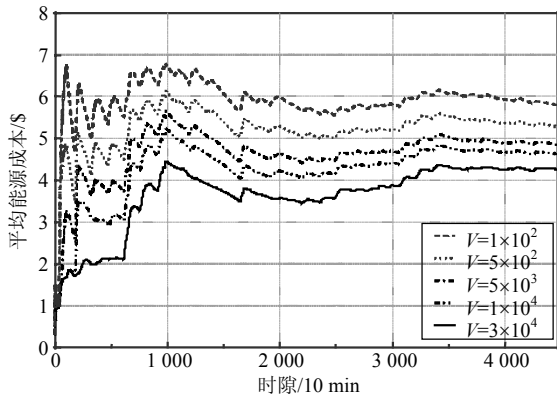


图3 V和平均成本关系

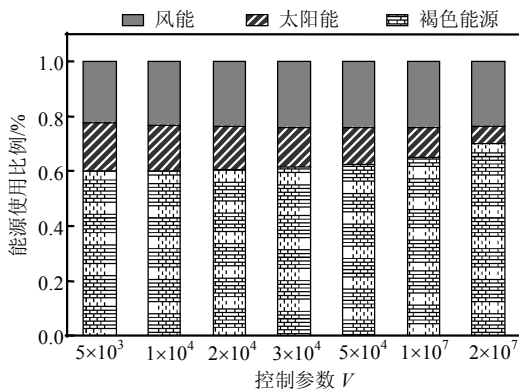


图4 V和能源使用比例关系

本文还分析了不同控制参数 V 下,数据中心各种能源的比例变化情况,如图4所示,其中碳排放约束 $H=55\text{ kg}$ 。由图4可见,当 V 增大时,褐色能源的使用比例逐渐增加,而太阳能使用比例则逐渐减少。原因在于 V 增加将会使式(16)中成本值优化比重增加,

意味着更大程度降低成本,从而促使选择价格更低的能源。同时,风能使用比例一直高于太阳能,原因在于风能比太阳能拥有具有更低的碳排放速率和价格,因而会被优先考虑。由此可知,如果不考虑风能和太阳能的部署因素,仅从价格和碳排放因素上看,风能将具有更好的竞争力。

2) 碳排放约束对能源成本的影响

当碳排放约束分别 50 kg , 60 kg , 70 kg , 80 kg 时,碳排放约束和能源成本关系如图5所示。由图5可见,随着平均碳排放约束的增大,算法将获得更低的成本,表明当放松碳排放约束时,数据中心运营成本将更低,与现实相符。但成本降低意味着温室气体排放增加,对气候环境影响愈加严重。

另外,不同碳排放约束下各种能源的使用比例如图6所示,其中 $V=30\text{ 000}$ 。由图6可见,随着碳排放约束增加,绿色能源使用比例逐渐减少,其中当碳排放约束为 80 kg 时,太阳能使用量已经接近为0。原因在于随着碳排放约束放松,算法将更多从价格上考虑能源选择,因此会优先选择低价的褐色能源、风能,而由于太阳能价格最高,所以用量最低。当碳排放约束为 85 kg 时,已经接近碳排放基准上界 89.05 kg ,说明此时碳排放对成本影响已经非常低。

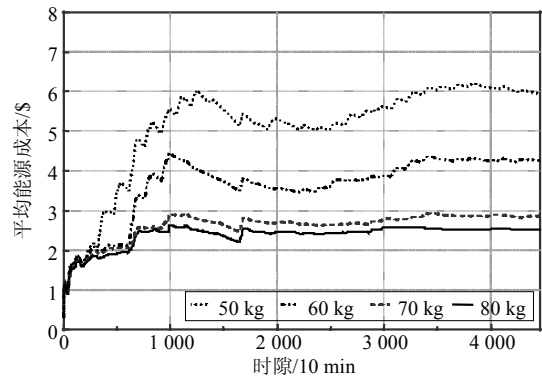


图5 碳排放约束和平均成本关系

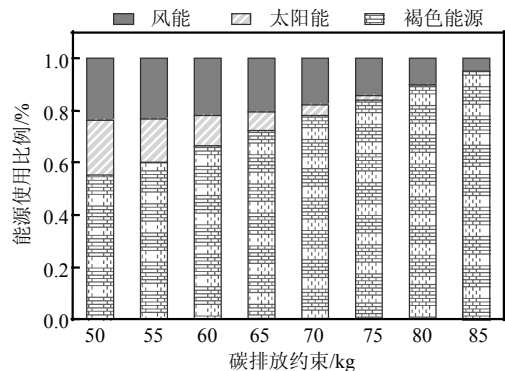


图6 碳排放约束和能源使用比例关系

3) 与相关算法比较

本文算法与Max-Green算法的比较如图7所示,此时碳排放约束 $H=50$ kg,控制参数 $V=10\ 000$ 。如前所述,Max-Green算法主要思想是实现数据中心最大绿色化^[38],其获得的碳排放约束下界为46.73 kg。在图7中,本文算法获得平均成本值为5.84,而Max-Green算法获得成本值为6.89。可见,在碳排放约束减少了6.8%的情况下,本文CACM算法可节省约13.8%的能耗成本,性能良好。同时本文算法可以根据不同碳排放约束获得其最优成本值,比Max-Green更为灵活。

本文算法与3.1节的离线最优算法比较如图8所示,此时碳排放约束 $H=70$ kg。由图8可见,随着控制参数 V 的增加,本文算法获得成本值会逐渐逼近离线最优成本,表示通过调节控制参数 V ,本文算法可以达到离线最优解。

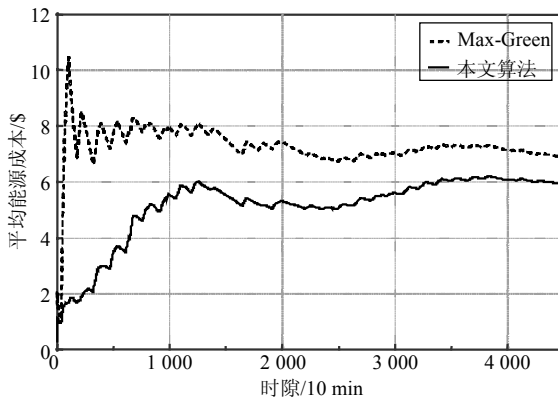


图7 本文算法与Max-Green算法对比

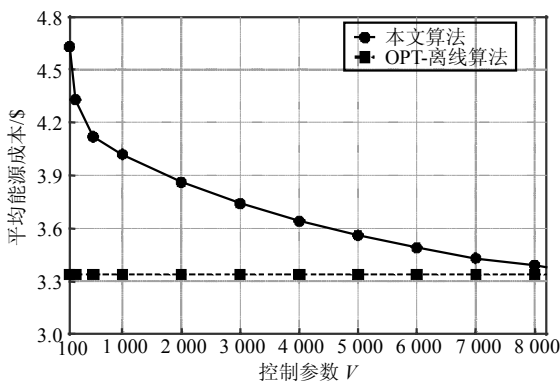


图8 本文算法和离线最优算法对比

5 结束语

本文针对满足不同类型负载的SLA和长期碳排放约束下,优化数据中心的能源成本问题,提出了一个碳感知的在线控制CACM算法。该算法不需要预测未来信息,仅靠当前系统状态信息作出决策,并可获得接近离线最优解的近似解。通过基于真实

数据的大量仿真实验验证了本文的CACM算法,结果表明CACM算法具有良好的性能,可以获得更优的能源成本。下一步工作计划将该算法推广到跨区域云数据中心能源成本优化研究,以及考虑利用能源存储的方式来降低数据中心的成本。

参 考 文 献

- [1] QURESHI A, WEBER R, BALAKRISHNAN H, et al. Cutting the electric bill for internet-scale systems[J]. ACM SIGCOMM Computer Communication Review, 2009, 39(4): 123-134.
- [2] KONG F, LIU X. A survey on green-energy-aware power management for datacenters[J]. ACM Computing Surveys, 2014, 47(2): 1-38.
- [3] GARTNER Inc. Gartner estimates ICT industry accounts for 2 percent of global CO₂ emissions[EB/OL]. [2016-10-20]. https://www.researchgate.net/publication/215697562_Gartner_Estimates_ICT_Industry_Accounts_for_2_Percent_of_Global_CO2_Emissions.
- [4] Google. Renewable energy[EB/OL]. [2016-10-20]. <http://www.google.com/green/energy/>.
- [5] Apple. Apple and the environment[EB/OL]. [2016-11-20]. <http://www.apple.com/environment/renewable-energy>.
- [6] RICH MILLER. 2008. Microsoft to use solar panels in new data center[EB/OL]. [2016-11-20]. <http://www.datacenterknowledge.com/archives/2008/09/24/microsoft-us-es-solar-panels-in-new-data-center/>
- [7] GUO Y, GONG Y, FANG Y, et al. Energy and network aware workload management for sustainable data centers with thermal storage[J]. IEEE Transactions on Parallel & Distributed Systems, 2014, 25(8): 2030-2042.
- [8] DENG W, LIU F, JIN H, et al. Harnessing renewable energy in cloud datacenters: Opportunities and challenges[J]. IEEE Network, 2014, 28(1): 48-55.
- [9] 邓维, 刘方明, 金海, 等. 云计算数据中心的新能源应用: 研究现状与趋势[J]. 计算机学报, 2013, 36(3): 582-598. DENG Wei, LIU Fang-ming, JIN Hai, et al. Leveraging renewable energy in cloud computing datacenters: State of the art and future research[J]. Chinese Journal of Computers, 2013, 36(3): 582-588.
- [10] RAO L, LIU X, XIE L, et al. Minimizing electricity cost: Optimization of distributed internet data centers in a multi-electricity-market environment[C]//Conference on Information Communications. San Diego: IEEE, 2010.
- [11] RAO L, LIU X, ILIC M D, et al. Distributed coordination of internet data centers under multiregional electricity markets[J]. Proceedings of the IEEE, 2012, 100(1): 269-282.
- [12] LIN M, LIU Z, WIERMAN A, et al. Online algorithms for geographical load balancing[C]//International Green Computing Conference. Washington DC, USA: IEEE, 2012.
- [13] LIU Z, LIN M, WIERMAN A, et al. Greening geographical load balancing[J]. IEEE/ACM Transactions on Networking, 2015, 23(2): 657-671.

- [14] YAO Y, HUANG L, SHARMA A, et al. Power cost reduction in distributed data centers: a two-time-scale approach for delay tolerant workloads[J]. IEEE Transactions on Parallel & Distributed Systems, 2014, 25(1): 200-211.
- [15] GOIRI, LE K, NGUYEN T D, et al. GreenHadoop: Leveraging green energy in data-processing frameworks[C]//ACM European Conference on Computer Systems. Bern: ACM, 2012: 57-70.
- [16] GOIRI I, LE K, HAQUE M E, et al. GreenSlot: Scheduling energy consumption in green datacenters[C]//Conference on High Performance Computing Networking, Storage and Analysis, SC 2011. Seattle, WA, USA: [s.n.], 2011.
- [17] 窦晖, 齐勇, 王培健, 等. 一种最小化绿色数据中心电费的负载调度算法[J]. 软件学报, 2014, 25(7): 1448-1458.
- DOU Hui, QI Yong, WANG Pei-jian, et al. Workload scheduling algorithm for minimizing electricity bills of green data centers[J]. Journal of Software, 2014, 25(7): 1448-1458
- [18] AKSANLI B, VENKATESH J, ZHANG L, et al. Utilizing green energy prediction to schedule mixed batch and service jobs in data centers[J]. ACM Sigops Operating Systems Review, 2012, 45(3): 53-57.
- [19] LIU Z, CHEN Y, BASH C, et al. Renewable and cooling aware workload management for sustainable data centers[J]. ACM SIGMETRICS Performance Evaluation Review, 2012, 40(1): 175-186.
- [20] KRIOUKOV A, ALSPAUGH S, MOHAN P, et al. Design and evaluation of an energy agile computing cluster[R]. Berkeley, UC: [s.n.], 2012.
- [21] ZHANG Y, WANG Y, WANG X. GreenWare: Greening cloud-scale data centers to maximize the use of renewable energy[C]//Proceedings of the 12th International Middleware Conference. Lisbon: ACM, 2011: 140-159.
- [22] REN S, HE Y. COCA: Online distributed resource management for cost minimization and carbon neutrality in data centers[C]//International Conference on High Performance Computing, Networking, Storage and Analysis. Denver: IEEE, 2013.
- [23] LIN M, WIERMAN A, ANDREW L L H, et al. Dynamic right-sizing for power-proportional data centers[J]. IEEE/ACM Transactions on Networking, 2011, 21(5): 1098-1106.
- [24] DEAN J, GHEMAWAT S. MapReduce: Simplified data processing on large clusters[J]. In Proceedings of Operating Systems Design and Implementation (OSDI), 2004, 51(1): 107-113.
- [25] CHEN R, YANG M, WENG X, et al. Improving large graph processing on partitioned graphs in the cloud[C]//ACM Symposium on Cloud Computing. San Jose: ACM, 2012: 1-13.
- [26] GUO Y, GONG Y, FANG Y, et al. Optimal power and workload management for green data centers with thermal storage[C]//GLOBECOM 2013 IEEE Global Communications Conference. Atlanta: IEEE, 2013.
- [27] CHOU W. Queueing systems, Volume II: Computer applications-Leonard Kleinrock[J]. IEEE Transactions on Communications, 1977, 25(1): 180.
- [28] RICHARD J. KOMP. Practical photovoltaics: Electricity from solar cells[M]. New York, USA: Aatec Publications, 1990.
- [29] GIPE P. Wind power: Renewable energy for home, farm, and business[M]. Claremont, USA: Chelsea Green Publishing Company, 2004.
- [30] DENG X, WU D, SHEN J, et al. Eco-aware online power management and load scheduling for green cloud datacenters[J]. IEEE Systems Journal, 2014, 10: 1-10.
- [31] ZHOU Z, LIU F, XU Y, et al. Carbon-aware load balancing for geo-distributed cloud services[C]//International Symposium on Modeling, Analysis & Simulation of Computer and Telecommunication Systems. San Francisco: IEEE, 2013: 232-241.
- [32] NEELY M J. Stochastic network optimization with application to communication and queueing systems[J]. Synthesis Lectures on Communication Networks, 2010, 3(1): 211.
- [33] CORPORATION I. IBM ILOG CPLEX optimizer[M]. Amon City, US: IBM Corporation, 2010.
- [34] NREL. NREL: Measurement and instrumentation data center[EB/OL]. [2016-10-20]. <http://www.nrel.gov/midc/>.
- [35] REN C, WANG D, URGANONKAR B, et al. Carbon-aware energy capacity planning for datacenters[C]//IEEE International Symposium on Modeling. Washington, DC: IEEE, 2012: 391-400.
- [36] STEWART C, SHEN K. Some joules are more precious than others: Managing renewable energy in the datacenter[C]//The Datacenter in Workshop on Hotpower. Montana, USA: [s.n.], 2009: 15-19.
- [37] NYISO. Market & operational data[EB/OL]. [2016-11-20]. <http://www.nyiso.com/>.
- [38] URDANETA G, PIERRE G, VAN STEEN M. Wikipedia workload analysis for decentralized hosting[J]. Elsevier Computer Networks, 2009, 53(11): 1830-1845.
- [39] LIU Z, LIN M, WIERMAN A, et al. Geographical load balancing with renewables[J]. ACM SIGMETRICS Performance Evaluation Review, 2011, 39(3): 62-66.