CN 11-2223/N J Tsinghu

# 分布式数据中心信息能量协同优化策略

刘迪<sup>1</sup>, 曹军威<sup>2</sup>, 刘明爽<sup>3</sup>

(1.清华大学自动化系,北京市 海淀区 100084;
2.清华大学北京信息科学与技术国家研究中心,北京市 海淀区 100084;
3. 深圳市腾讯计算机系统有限公司,深圳市 南山区 518057)

摘要:随着数据中心规模的不断扩大,其能耗巨大的问题也 日益突出。分布式数据中心既可以通过计算任务在多个数据 中心间的分配实现功率的转移,也可以通过单个数据中心的 功率控制实现功耗和计算时延的均衡。两种优化手段相互耦 合,且面临着来自于信息层和能量层的多元不确定性的影响, 需要快速可靠的控制手段实现数据中心信息层和能量层的协 同优化。基于此,本文首先构建了分布式数据中心协同优化 调节架构,并基于动态微分方程,分析了多数据中心计算任 务分配与单数据中心功率优化的动态特性,构建了耦合优化 问题的统一调节模型。综合考虑系统运营成本及计算时延构 建目标函数,引入最优控制理论对该问题求解,实现数据中 心信息能量的秒级协同优化控制。仿真结果表明,相较于分 钟级的控制,基于本文所提算法的快速控制能够较好的追踪 可再生能源出力以及计算任务的波动,从而有效提升系统的 经济效益及可再生能源就地消纳率。

关键词:分布式数据中心、微分方程、协同优化、可再生能 源,最优控制

# Collaborative optimization strategy of information and energy for distributed data centers

#### LIU Di<sup>1</sup>, CAO Junwei<sup>2</sup>, LIU Mingshuang<sup>3</sup>

(1. Department of Automation, Tsinghua University, Haidian District, Beijing 100084, China;

2. Beijing National Research Center for Information Science and Technology, Tsinghua University, Haidian District, Beijing 100084, China

3. Shenzhen Tencent Computer System Co., Ltd., Nanshan District, Shenzhen 518057, China)

**Abstract:** With the continuous expansion of the scale of data centers, the problem of huge energy consumption has become increasingly prominent. Distributed data centers can not only realize power transfer through the distribution of computing tasks among multiple data centers, but also realize the balance of power consumption and computing delay through power control

of a single data center. The two optimization methods are coupled with each other and face the influence of multiple uncertainties from the information layer and the energy layer. Therefore, a fast and reliable control method is required to realize the collaborative optimization of the information layer and the energy layer of the data center. Based on this, a distributed data center collaborative optimization architecture was first constructed in this paper. Then the dynamic characteristics of multi data center computing task allocation and single data center power optimization are analyzed based on the dynamic differential equation, and a unified adjustment model of coupling optimization problem is constructed. Considering the system operating cost and computing delay to construct the objective function, the optimal control theory is introduced to solve this problem and realize the second-level collaborative optimal control of the information energy of the data center. Simulation results show that, compared with the minute-level control, the high frequency control based on the proposed algorithm can better track the fluctuation of renewable energy output and calculation tasks, so as to effectively improve the economic benefits of the system and the local consumption rate of renewable energy.

**Key words:** Distributed data center, differential equation, collaborative optimization, renewable energy, optimal control

随着数字化社会的发展,数据中心的规模在不断扩大,其能耗问题也日益凸显。在最近的几年间, 我国数据中心的用电量每年的增幅超过 10%,2020 年突破 2000 亿 kWh,占全社会总用电量的 2.7%。 根据国网能源研究院的预测,到 2030 年,全国数据 中心的总用电量会在 2020 年的基础上翻倍,占全社 会用电量的比重也将上升至 3.7%。

与此同时,我国提出"碳中和、碳达峰"的目标,促使电力系统向着以清洁能源为主的结构转变。 但可再生能源出力的波动性和不可调节[1]的特性使

收稿日期: yyyy-mm-dd 基金项目:腾讯高校合作犀牛鸟基础平台技术专项 作者简介:刘迪(1990-),男,博士后。 通信作者:曹军威,研究员,E-mail: jcao@tsinghua.edu.cn 高比例可再生能源电网的供给侧灵活性大大降低<sup>[2]</sup>, 需要充分利用储能<sup>[3]</sup>、可调节负荷<sup>[4]</sup>等需求侧可调节 资源提升电网灵活性,消纳可再生能源。数据中心 作为高能耗负荷之一,其目前的能量来源还是以传 统化石能源为主,2018 年,可再生能源仅仅供应了 23%的数据中心能耗,低于可再生能源在中国全社会 电力消费量的占比 26.5%<sup>[5]</sup>。因此,高能耗和能源供 给中可再生能源占比低就成了数据中心运营中面临 的两大问题。

目前的研究中,主要从数据中心本身的能效提 升<sup>[6]</sup>以及数据中心与电网协同优化<sup>[7]</sup>两方面展开。针 对高能耗问题,除了从数据中心建设本身提升数据 中心能源利用率外,在满足时延要求的前提下适当 降低数据中心的服务率,能够在一定程度上降低数 据中心的能耗<sup>[8][9]</sup>。针对能源供给中可再生能源占比 低的问题,可以通过计算任务在不同数据中心之间 的灵活分配实现负荷的迁移<sup>[10]</sup>,以更多的消纳可再 生能源。具体而言,可以通过直接负荷控制、电价 信号等手段<sup>[11]</sup>,引导数据中心参与电网需求响应, 以平抑新能源波动、缓解电网阻塞,提升系统运行 效率<sup>[12]</sup>。

文献[13]对混合供电的数据中心的能耗优化问 题进行了综述,分别从数据中心本身能效的提升以 及多个数据中心负载均衡等角度出发,综述了目前 的最新研究成果,并对数据中心能耗与可再生能源 出力的联合调节提出了研究展望,但并未深入的研 究。文献[14]提出了一种多数据中心流量分配的实时 算法,通过计算任务的优化分配,提升数据中心的 长期效益。文献[15]在此基础上,进一步考虑了时延 等因素,提出了分布式的多数据中心流量分配实时 算法,以最小化多个数据中心的长期运营成本。文 献[16]建立了数据网络与电力网络混合运行模型,将 数据中心任务分配纳入电网经济运行的优化调度 中。

在上述研究中,通过不同数据中心之间的计算 任务调度,实现数据中心功耗的再分配,提升系统 的总体效益。但其前提假设是在特定的计算任务下 数据中心的功耗是确定的。事实上,根据计算任务 时延的需求,可以通过服务器的开关<sup>[17]</sup>以及 CPU 频 率的调整<sup>[18]</sup>动态调整数据中心的使用率,进一步提 升系统总体效率。数据中心的能耗优化问题,就转 变为两个子问题:(1)考虑能源层经济成本的信息 层计算任务分配;(2)考虑信息层计算时延的能源 层数据中心能耗的确定。

上述两个子问题对于系统总体效益的影响是相 互耦合的<sup>[7]</sup>。针对上述优化问题,文献[17]构建了双 层优化模型,在日前根据预测的计算任务量及时延 需求,计算出数据中心可调节功率范围并参与电力 市场出清,日内根据出清结果调节数据中心的任务 量及功率。文献[18]也构建了电网和和信息双层优化 架构,通过两层间的交互优化,保证计算效率的同 时,降低电网阻塞率,同样需要在日前向电网提供 数据中心可调度的功率范围。文献[19]则是在计算任 务预测的基础上,通过两层之间的多次迭代,逐渐 逼近最优决策。文献[20]综合考虑数据中心和电动汽 车的调度,将求解问题转化为随机规划问题并基于 ADMM 算法求解该问题。

上述研究中,均构建了能量和信息的双层优化 调节架构,通过数据预测或信息交互迭代实现数据 中心和电网的联合优化,实现了小时级/分钟级的优 化调节。但数据中心计算任务存在较大的波动性<sup>[24]</sup>, 且常常会面临难以预测的流量突发现象<sup>[10]</sup>,同时, 可再生能源的出力也具有很大的波动性,信息层和 能量层的双重不确定性对于控制的反馈速度提出了 更高的要求。相比于迭代的求解算法,最优控制能 够实现毫秒级的反馈,并且已经在火箭运行姿态调 整<sup>[21]</sup>,电力系统稳定控制<sup>[22]</sup>,燃料电池的能量管理<sup>[23]</sup> 等领域取得了较好的应用效果。

因此,为了提升分布式数据中心控制的速度以 更好的应对突发计算任务以及可再生能源出力带来 的波动,本文引入最优控制理论,针对目前研究中 多数据中心计算任务的分配与单数据中心功率控制 协同优化的问题,基于动态微分方程,构建多数据 中心计算任务分配及数据中心功率优化动态控制统 一模型,实现任务分配与功率控制的同步优化。进 而,采用最优控制的极小值原理求解对控制问题求 解,有效提升了系统控制的速度,保证系统能够及 时随外界环境变化做出调整,使系统总能够根据外 界环境的波动修正最优控制决策,并对系统控制速 度与效益的关系进行了分析。此外,考虑到数据处 理的服务质量(Quality of Service, QoS)与经济成本是相互矛盾的,本文在调节目标中兼顾了延迟成本与经济效益,并探讨了 QoS 与经济效益之间的关系。

# 分布式数据中心架构及信息能量模型 1.1 分布式数据中心总体架构

分布式数据中心分布在电网的不同节点,一方 面,数据中心可以依托电网供电,另一方面,数据 中心可以依托附近可再生能源供电<sup>[25]</sup>。在信息层, 各个数据中心互联互通,来自于前端的计算任务可 以自由分配,分布式数据中心的信息和能源架构如 图1所示:





如图 1 所示,在每个时刻,有 k 个前端申请计算 任务,每个前端申请的计算任务量为 l<sub>i</sub>(t),所有计 算任务可以分配给 n 个分布式数据中心协同完成, 每个数据中心承担 L<sub>DC,i</sub>(t)的计算任务。各个数据中 心之间在信息层相互连通,但由于地理上的分布特 性,其通过不同的主网节点接入电网,从而每一个 数据中心的实时节点电价各不相同<sup>[26]</sup>。数据中心除 了能从电网购电外,还可以依托本地可再生能源功 能,充分利用可再生能源"剩余"电量,以减少弃 风弃光现象的发生。由于各地资源禀赋及负荷特性 不同,能够给数据中心供能的可再生能源的容量及 时域特性也各不相同。

# 1.2 分布式数据中心信息能量模型

#### 1.2.1 数据中心能耗控制模型

数据中心的能耗与承担的计算任务量相关,通 过数据中心服务器的启停<sup>[17]</sup>以及 CPU 频率的调节 <sup>[18]</sup>,可以控制数据中心的服务率,使数据中心不总 运行在满功率的状态下,以提升数据中心的能效。

数据中心的能耗主要包括服务器能耗以及支持服务器正常运行的散热等能耗构成<sup>[27]</sup>。其中,数据中心服务器的能耗主要为 CPU 的能耗,可以用如下模型表示<sup>[6][28]</sup>:

$$P_{\text{server},i,j}(t) = c_{i,j,0} + A_{i,j}E_{i,j}V_{i,j}^{2}(t)f_{i,j}(t)$$
(1)

上式中数据中心的服务器功率主要包括静态功率和动态功率,  $c_{i,j,0}$ 为第 i 个数据中心第 j 个服务器 CPU 的静态功率, 如果硬件环境不变,则其为固定 常数。服务器的动态功率则由电容  $E_{i,j}$ ,时钟频率  $f_{i,j}(t)$ , CPU 电压  $V_{i,j}(t)$ 共同构成,其能耗与每个时 钟周期内的开关次数  $A_{i,j}$  正相关。其中  $A_{i,j}$ 和  $E_{i,j}$ 在相 同的硬件环境下也为固定常数。

根据数据中心的典型电能利用效率(Power Usage Effectiveness, PUE)<sup>[29]</sup>,可以得到数据中心服务器功率与总功率的关系:

$$P_{DC,i}(t) = \sum_{j=1}^{M_i} P_{server,i,j}(t) U_i$$
(2)

其中, $U_i$ 为第*i*个数据中心的 PUE 值。如前文 所述,可以通过控制开启服务器的数量<sup>[17]</sup>或调节 CPU 频率<sup>[18]</sup>实现数据中心的功率控制,同时,相对 于整个数据中心而言,功率可调节的颗粒度很小, 因此认为数据中心的功率是连续可调的,可以定义 数据中心使用率控制信号为 $u_{P_i}(t) \in [0,1]$ ,以控制数 据中心功率的变化。

对于一个数据中心而言,如果没有新到计算任 务,随着数据中心存量任务的处理,数据中心的所 需处理的计算负载在逐渐下降,则维持同样计算时 延所需的功率也在逐渐下降。假设数据中心的功率 能够追踪计算负载的变化,则在固定计算时延下, 数据中心功率与负载之间存在正相关的关系,如下 所示:

$$P_{DC,i}(t) = k_i L_{DC,i}(t)$$
(3)

其中, $k_i$ 是默认情况下,数据中心功率 $P_{DC,i}(t)$ 与 计算任务 $L_{DC,i}(t)$ 间的比例系数,使数据中心的功率 能随计算任务量的下降而降低。 同时,数据中心对任务的处理速度也与功率相关,则剩余计算任务的变化量为:

$$dL_{DC,i}(t) = -v_i P_{DC,i}(t) \tag{4}$$

其中, $v_i$ 是计算任务变化量与数据中心功率的比例系数,计算中心功率 $P_{DC_i}(t)$ 越大,数据处理速度越快,则单位时间被处理完成的计算任务 $dL_{DC_i}(t)$ 越多。

结合(3)和(4),在特定的控制信号下,每个数据 中心的功率随时间的动态变化可以表示为如下形 式:

$$dP_{DC,i}(t) = -k_{i}v_{i}P_{DC,i}(t) + \beta_{P,i}u_{P,i}(t)$$
(5)

其中,  $\alpha_{P,i}$  为时间常数,  $\beta_{P,i}$  为最大可到达计算 任务,  $\beta_{P,i}u_{P,i}(t)$  表示 t 时隙中第 i 个数据中心新增的 计算任务。

数据中心功率的变化,会影响数据中心能量层 的成本。每个数据中心的能量来源一方面来自本地 的分布式可再生能源,一方面来自于电网,且当本 地可再生能源过多时,可以出售给电网,获得一定 的售电收益,因此,所有数据中心*t*时刻的能源成本 如下:

$$C(t) = \sum_{i=1}^{n} \left( P_{B,i}(t) R_{B,i}(t) \Delta t - P_{S,i}(t) R_{S,i}(t) \Delta t \right)$$
  
s.t.  $P_{PV,i}(t) + P_{B,i}(t) - P_{S,i}(t) = P_{DC,i}(t)$  (6)

上式表示由 n 个数据中心组成的数据中心系统 的能源成本 C(t),当 t 时隙数据中心 i 的本地可再生 能源发电量不能满足数据中心能耗时,其会向电网 购买电量  $P_{B,i}(t)$ ,其接入的主网处的购电电价为  $R_{B,i}(t)$ 。当 t 时隙数据中心 i 的本地可再生能源发电 量超过数据中心能耗需求时,剩余部分电量会出售 给电网,出售电量用  $P_{S,i}(t)$ 表示,t 时隙第 i 个数据 中心所处电网节点的可再生能源的上网价格为  $R_{S,i}(t)$ 。同时,每个数据中心可再生能源出力、数据 中心能耗以及与电网之间的交互功率应实时平衡, 即满足功率平衡约束。

#### 1.2.2 数据中心计算任务分配模型

如图1所示,在每个时隙*t*,数据中心根据*k*个 前端的计算请求分配每个数据中心计算的任务量, 因此定义数据中心计算任务的分配控制信号为  $u_{L,i}(t) \in [0,1]$ ,则在分配控制信号 $u_{L,i}(t)$ 下,每个数据中心每个时隙承担的计算任务量的动态变化如下式所示:

$$dL_{DC,i}(t) = -k_i v_i L_{DC,i}(t) + \beta_{L,i} u_{L,i}(t)$$
(7)

其中,  $u_{L,i}(t)$  为分配给第*i*个数据中心的任务的 控制量,  $\beta_{L,i}$  为可以分配给第*i*个数据中心的最大的 任务量,  $\beta_{L,i}u_{L,i}(t)$  表示 *t* 时隙中分配给第*i*个数据中 的计算任务量。为了保证在前端的计算请求均能够 被数据中心处理,设置约束如下:

$$\int_{0}^{T} \sum_{i=1}^{n} L_{DC,i}(t) dt \ge \int_{0}^{T} \sum_{j=1}^{k} l_{j}(t) dt$$
(8)

上式保证了所考虑的时间段 T 内,每个数据中 心承担的计算任务量  $L_{DC,i}(t)$  的总和始终大于等于来 自于前端计算请求  $l_i(t)$  的总和。

#### 1.2.3 考虑延迟的信息能量耦合模型

就信息层而言,数据处理的延迟是数据中心服务质量的重要指标,数据中心计算任务的平均延迟 又与数据中心承担的计算任务量以及数据中心的功 率有关。在目前的研究中,常用 M/M/n 排队模型计 算数据中心的平均时延情况<sup>[17]-[19][30]</sup>:

$$D_{i}^{P} = \frac{1}{\sum_{j=1}^{M_{i}} \mu_{i,j}(t) - \sum_{j=1}^{M_{i}} L_{DC,i,j}(t)}$$
(9)

其中, $D_i$ 为第*i*个数据中心的平均时延, $\mu_{i,j}(t)$ 为数据中心服务器的服务率,服务率越高,数据中 心的延迟越低, $M_i$ 为第*i*个数据中心服务器的数量。 数据中心功率的调节手段有开关服务器数量<sup>[17]</sup>以及 调节服务器的 CPU 频率<sup>[18]</sup>,关闭服务器会增加每个 服务器分担的网络负载,在计算延迟时,可以认为 是服务器数量不变,每台服务器的服务率 $\mu_{i,j}(t)$ 下 降,从而在计算时总有 $M_i$ 台虚拟服务器,保持服务 器数量不变,以方便延迟的计算以及后续的优化。 需要说明的是,实际环境中的任务分配更为复杂, 本文参考文献<sup>[17]-[19]</sup>等的研究成果,假设数据中心内 部已存在某种任务分配方式,将数据中心看做一个 整体,用该简化的映射关系表示数据中心功率与时 延的外部特性。

此外,由于网络阻塞的原因,计算任务从前端 分配至每个数据中心,也会带来时延,同样可以根 据 M/M/1 排队模型,假设计算任务的数据大小呈指 数分布,考虑了网络阻塞后,(9)式可以被改写为如下形式<sup>[31]</sup>:

$$D_{i} = \frac{\varpi}{\sum_{j=1}^{M_{i}} \mu_{i,j}(t) - \varpi \sum_{j=1}^{M_{i}} L_{DC,i,j}(t)}$$
(10)

其中, *σ* 是网络时延参数,取决于网络没有阻 塞时的网络传输时延。

在能量层面,数据中心的服务率 $\mu_{i,j}(t)$ 与 CPU 频率正相关<sup>[32]</sup>,即 $\mu_{i,j}(t) \propto f_{i,j}(t)$ ,又由式(1)可知, CPU 频率与功耗近似线性关系,因此数据中心服务 率与功耗之间的关系可由下式近似表示:

$$\sum_{j=1}^{M_{i}} \mu_{i,j}(t) = \varphi_{i} \sum_{j=1}^{M_{i}} P_{server,i,j}(t)$$
(11)

其中, φ<sub>i</sub>为数据中心服务器功率与服务率的转 换系数。联立式(2)(10)(11),可得数据中心功耗与计 算延迟之间的关系:

$$D_{i} = \frac{\varpi}{\frac{\varphi_{i}}{U_{i}} P_{DC,i}\left(t\right) - \varpi \sum_{j=1}^{M_{i}} L_{DC,i,j}\left(t\right)}$$
(12)

综上,本文构建了能量层的数据中心的能耗模型、信息层的计算任务分配模型以及数据中心功耗 与平均计算延迟的能量信息耦合模型。

### 2 问题构建及求解

#### 2.1 问题构建

从前述的模型中可以看出,在每个时隙,数据 中心需要决策的变量为分配的计算任务量以及数据 中心的功率。对于*t* ∈ [0,*T*],数据中心任务量分配以 及功率调节的控制系统可以定义为:

$$\begin{cases} \dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{x}(0) = \mathbf{x}_0 \end{cases}$$
(13)

其中, **A** 和 **B** 均为系数矩阵,  $\mathbf{x}_0$  为 0 时隙的系 统初始状态,  $\mathbf{x}(t) \ge n$  个数据中心的状态变量:

$$\mathbf{x}(t) = \begin{bmatrix} L_{\text{DC},1}(t), P_{\text{DC},1}(t), L_{\text{DC},2}(t), \\ P_{\text{DC},2}(t), \dots, L_{\text{DC},n}(t), P_{\text{DC},n}(t) \end{bmatrix}^{\text{T}}$$
(14)

右上角标"T"表示矩阵的转置, L<sub>DC,i</sub>(t)是第 i 个数据中心 t 时隙的计算任务量, P<sub>DC,i</sub>(t)是第 i 个数 据中心 t 时隙的功率。u(t)是 n 个数据中心的控制输 入:

$$\mathbf{u}(t) = \begin{bmatrix} u_{L,1}(t), u_{P,1}(t), u_{L,2}(t), \\ u_{P,2}(t), \dots, u_{L,n}(t), u_{P,n}(t) \end{bmatrix}^{\mathrm{T}}$$
(15)

其中 $u_{L,i}(t) \in [0,1]$ , 和 $u_{P,i}(t) \in [0,1]$ 分别为t时隙 第i个数据中心计算任务的控制量信号和功率的控 制信号。根据式(5)(7)可得对称系数矩阵 A 和 B 分别 为:

$$\mathbf{A} = \begin{bmatrix} -k_{1}v_{1} & \cdots & 0 \\ & -k_{1}v_{1} & & \\ \vdots & \ddots & \vdots \\ & & -k_{n}v_{n} \\ 0 & \cdots & -k_{n}v_{n} \end{bmatrix}$$
(16)  
$$\mathbf{B} = \begin{bmatrix} \beta_{L,1} & \cdots & 0 \\ & \beta_{P,1} & & \\ \vdots & \ddots & \vdots \\ & & & \beta_{L,n} \\ 0 & \cdots & & \beta_{P,n} \end{bmatrix}$$
(17)

数据中心的能耗成本及功率平衡约束如式(6)所 示,除此之外,数据中心*t*时隙的购电电量应大于等 于 0,即  $P_{B,i}(t) \ge 0$ ,数据中心*t*时隙出售给电网的电 量 应 介 于 可 再 生 能 源 出 力 与 0 之 间 ,即  $P_{PV,i}(t) \ge P_{S,i}(t) \ge 0$ 。为了促进可再生能源的就地消 纳,可再生能源的上网价格会低于购电价格<sup>[33]</sup>,因 此,数据中心会优先使用可再生能源发电为数据中 心供电,则:

$$P_{S,i}(t) = \begin{cases} P_{PV,i}(t) - P_{DC,i}(t), P_{PV,i}(t) \ge P_{DC,i}(t) \\ 0, P_{PV,i}(t) < P_{DC,i}(t) \end{cases}$$
(18)

$$P_{B,i}(t) = \begin{cases} P_{DC,i}(t) - P_{PV,i}(t), P_{PV,i}(t) < P_{DC,i}(t) \\ 0, P_{PV,i}(t) \ge P_{DC,i}(t) \end{cases}$$
(19)

从而,式(6)中的数据中心在 *t* 时隙的单位成本可 以改写为如下形式:

$$C(t) = \sum_{i=1}^{n} \left( P_{DC,i}(t) - P_{PV,i}(t) \right) R_i(t) \Delta t \qquad (20)$$

其中,

$$R_{i}(t) \begin{cases} R_{B,i}(t), P_{PV,i}(t) < P_{DC,i}(t) \\ R_{S,i}(t), P_{PV,i}(t) \ge P_{DC,i}(t) \end{cases}$$
(21)

除了经济成本之外,服务质量也是数据中心能 耗优化需要考虑的重要因素之一。因此,结合式(12), 定义目标函数如下:

$$\mathbf{L}\left[x(t)\right] = \sum_{i=1}^{n} \left(\gamma D_{i} + \left(P_{DC,i}(t) - P_{PV,i}(t)\right)R_{i}(t)\right)$$
(22)

其中 γ ≥ 0,为权重因子, γ 越大,表示目标对 延迟的要求越高。对于延迟而言,一方面希望实现 其与经济成本之间的均衡,另一方面其也应满足最 大延迟约束,结合式(12)可定义延迟约束:

$$D_i \le t_{\text{delay},i}, i=1,\dots,n \tag{23}$$

其中, t<sub>delay,i</sub> 即为第 *i* 个数据中心的延迟上限, 改 写上式可得:

$$\frac{\varphi_i}{U_i} P_{DC,i}(t) - \varpi L_{DC,i}(t) \ge \frac{\varpi}{t_{\text{delay},i}}, i = 1, ..., n$$
(24)

为了在控制中满足延迟约束,结合式(24)定义  $L_1(x(t)) = Sx(t) + D$ ,其中,

$$\mathbf{S} = \begin{bmatrix} \overline{\boldsymbol{\varpi}} & -\frac{\varphi_1}{U_1} & \cdots & 0 & 0\\ 0 & 0 & \vdots & \vdots\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \vdots & \vdots & 0 & 0\\ 0 & 0 & \cdots & \overline{\boldsymbol{\varpi}} & -\frac{\varphi_n}{U_n} \end{bmatrix}$$
(25)
$$\mathbf{D} = \begin{bmatrix} \overline{\boldsymbol{\varpi}} \\ t_{delay,1}, \dots, \overline{\boldsymbol{\varpi}} \\ t_{delay,n} \end{bmatrix}^{\mathrm{T}}$$
(26)

则不等式约束可以表示为:

$$\mathbf{L}_{1}[\mathbf{x}(t)] \prec 0 \tag{27}$$

符号"≺"表示矩阵 L<sub>1</sub>[x(t)]中的每一个元素均 小于等于 0。此外,式(12)所表示的系统应在考虑的 时间内,处理完所有用户的任务请求,即满足式(8) 所表示的约束,对其移项变形可得:

$$\int_{0}^{T} \left( \sum_{i=1}^{n} L_{DC,i}(t) - \frac{1}{T} \sum_{j=1}^{k} l_{j}(t) \right) dt \ge 0$$
 (28)

为了在控制中满足任务处理量的约束,根据式 (28)定义 L<sub>2</sub>:

$$L_{2}\left(\mathbf{x}(t)\right) = \frac{1}{T} \sum_{j=1}^{k} l_{j}\left(t\right) - \boldsymbol{\alpha}^{\mathrm{T}} \mathbf{x}(t)$$
(29)

其中, *α* = [1,0,1,0,...,1,0]<sup>T</sup>,则积分约束可以表 示为:

$$\int_0^T L_2\left(\mathbf{x}(t)\right) \le 0 \tag{30}$$

每个数据中心的任务分配量以及功率可以在 0 到最大值之间被调节。因此,定义每个数据中心任 务分配量和功率的容许控制集 u(t):

$$\Gamma = \left\{ \mathbf{u}(t) \middle| \mathbf{0} \prec \mathbf{u}(t) \prec \mathbf{1} \right\}$$
(31)

最终构建数据中心任务分配及功率确定的最优 控制问题如下:

$$\min_{\mathbf{u}(t)\in\Gamma} J\left(0, \mathbf{x}(0), \mathbf{u}(\Box)\right) = \int_{0}^{T} \mathbf{L}\left[\mathbf{x}(t)\right] dt$$
  
s.t. 
$$\begin{cases} \mathbf{L}_{1}\left[\mathbf{x}(t)\right] < 0, t \in [0, T] \\ \int_{0}^{T} L_{2}\left[\mathbf{x}(t)\right] dt \le 0 \end{cases}$$
 (32)

经过上述推导与转换后,数据中心计算任务分 配以及最佳功率的选择就被转化为一个最优控制问 题,而后就需要对该系统求解,以获得数据中心每 一时隙最优的控制策略。

#### 2.2 问题求解

针对式(13)的系统,式(32)的最优控制问题既有 不等式过程约束,也有不等式积分约束。为了在最 优控制的过程中满足约束,分别引入 Lagrange 乘子:  $\lambda(t)$ ,  $\eta(t)$ ,  $\zeta$ 。首先考虑不等式积分约束,定义 Hamilton 函数 *H*:

$$H[\mathbf{x}(t),\mathbf{u}(t),\lambda(t),\zeta,t] = \mathbf{L}[\mathbf{x}(t)] + \lambda^{\mathrm{T}}(t)[\mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)] + \zeta L_{2}[\mathbf{x}(t)]$$
(33)

其中, $\lambda(t)$ 为系统约束的参数, $\zeta$ 为不等式积分 约束项的参数。进而考虑不等式过程约束,定义 Lagrange 函数 $\hat{L}$ :

$$\hat{L}[\mathbf{x}(t),\mathbf{u}(t),\lambda(t),\eta(t),\zeta,t] = 
H[\mathbf{x}(t),\mathbf{u}(t),\lambda(t),t] + \boldsymbol{\eta}^{\mathrm{T}}(t)\mathbf{L}_{1}[\mathbf{x}(t)]$$
(34)

其中,  $\eta(t)$ 为不等式过程约束项的参数。

由最优控制理论的极小值原理<sup>[34]</sup>,式(32)中问题 的最优控制信号 u<sup>\*</sup>(t)和最优轨线,即每个数据中心 的负荷分配量与功率随时间最佳的状态变化曲线 x<sup>\*</sup>(t)应满足最优控制原理求解中所需的方程、等式 和不等式。 根据式(13)的形式,数据中心的最优计算任务分 配及最优功率选择需满足如下规范方程:

$$\dot{\mathbf{x}}^{*}(t) = \mathbf{A}\mathbf{x}^{*}(t) + \mathbf{B}\mathbf{u}^{*}(t)$$
(35)

λ(t)应满足如下规范方程:

$$\begin{split} \dot{\boldsymbol{\lambda}}(t) &= -\frac{\partial \hat{L}\left[\mathbf{x}(t), \mathbf{u}(t), \boldsymbol{\lambda}(t), \eta(t), \boldsymbol{\zeta}, t\right]}{\partial \dot{\mathbf{x}}^{*}(t)} = \\ \begin{bmatrix} \left(-\gamma \overline{\boldsymbol{\varpi}} \middle/ \left(\frac{\varphi_{1}}{U_{1}} P_{DC,1}^{*}(t) - \overline{\boldsymbol{\varpi}} L_{DC,1}^{*}(t)\right) + \eta_{1}(t)\right) + \boldsymbol{\zeta} \\ \frac{\varphi_{1}}{U_{1}} \left(\gamma \overline{\boldsymbol{\varpi}} \middle/ \left(\frac{\varphi_{1}}{U_{1}} P_{DC,1}^{*}(t) - \overline{\boldsymbol{\varpi}} L_{DC,1}^{*}(t)\right) + \eta_{1}(t)\right) - R_{1}(t) \\ \vdots \\ \left(-\gamma \overline{\boldsymbol{\varpi}} \middle/ \left(\frac{\varphi_{n}}{U_{n}} P_{DC,n}^{*}(t) - \overline{\boldsymbol{\varpi}} L_{DC,n}^{*}(t)\right) + \eta_{n}(t)\right) + \boldsymbol{\zeta} \\ \frac{\varphi_{n}}{U_{n}} \left(\gamma \overline{\boldsymbol{\varpi}} \middle/ \left(\frac{\varphi_{n}}{U_{n}} P_{DC,n}^{*}(t) - \overline{\boldsymbol{\varpi}} L_{DC,n}^{*}(t)\right) + \eta_{n}(t)\right) - R_{n}(t) \end{bmatrix} \end{split}$$
(36)

同时,在对数据中心控制过程中,需要获得开 始控制时数据中心及能源系统的初始状态,以及控 制结束时的系统状态,因此需要满足边值条件 **x**(0)=**x**<sub>0</sub>以及λ(*T*)=0。结合式(36)可得:

$$\lambda_{2i-1}(t) = \int_{T}^{t} e^{(t-\tau)k_{i}v_{i}} \left[ -\left(\frac{\gamma \varpi}{(\varphi_{i}/U_{i})P_{DC,i}^{*}(\tau) - \varpi L_{DC,i}^{*}(\tau)} + \eta_{i}(\tau)\right) + \zeta \right] d\tau$$

$$(37)$$

$$\lambda_{2i}(t) = \int_{T}^{t} e^{(t-\tau)k_{i}v_{i}} \left[ \frac{\varphi_{i}}{U_{i}} \left( \frac{\gamma \varpi}{(\varphi_{i}/U_{i})} P_{DC,i}^{*}(\tau) - \varpi L_{DC,i}^{*}(\tau) + \eta_{i}(\tau) \right) - R_{i}(\tau) \right] d\tau$$
(38)

为了保证数据中心控制的控制决策能使目标函 数能够达到极小值,满足如下极小值条件:

$$H\left[\mathbf{x}^{*}(t),\mathbf{u}^{*}(t),\lambda(t),\zeta,t\right] = \min_{\mathbf{u}(t)\in\Gamma} H\left[\mathbf{x}^{*}(t),\mathbf{u}(t),\lambda(t),\zeta,t\right]$$
(39)

即 $\lambda^{\mathrm{T}}(t)$ **Bu**<sup>\*</sup>(t) =  $\min_{\mathbf{u}(t)\in\Gamma}\lambda^{\mathrm{T}}(t)$ **Bu**(t),从而可以得到 使数据中心总体的成本和时延最低的控制决策如 下:

$$u_{i}^{*}(t) = \begin{cases} 1, \lambda_{i}(t) < 0\\ 0, \lambda_{i}(t) > 0 \end{cases}, i = 1, ..., 2n$$
(40)

经过上述推导,数据中心计算任务分配以及功 率控制被转化为最优控制中可以求解的形式,在系 统求解时应满足约束条件,如式(40)所示。

联立上述各式可得:

$$\begin{cases} \boldsymbol{\zeta} \ge 0 \\ \int_{0}^{T} L_{2} [\mathbf{x}(t)] dt \le 0 \\ \boldsymbol{\zeta} \int_{0}^{T} L_{2} [\mathbf{x}(t)] dt = 0 \\ \boldsymbol{\eta}(t) \succ 0 \\ \mathbf{L}_{1} [\mathbf{x}(t)] \prec 0 \\ \boldsymbol{\eta}^{\mathrm{T}}(t) \mathbf{L}_{1} [\mathbf{x}(t)] = 0 \end{cases}$$
(41)

通过 GEKKO 工具包<sup>[35]</sup>求解该等式/不等式方程 组,可得到问题(32)的最优控制信号 u<sup>\*</sup>(t)以及相对 应的负荷分配量以及每个数据中心功率随时间变化 的曲线 x<sup>\*</sup>(t)。

#### 3 仿真实验

#### 3.1 仿真环境设置

本文以 3 个数据中心为例进行仿真,数据中心的参数如表 1 所示:

表1数据中心仿真参数

	DC1	DC2	DC3
服务器数量(台)	2*105	2*105	2*105
服务器最大功率(W)	300	300	300
服务器满载处理速率(个/s)	30	30	30
$\varphi$	0.1	0.1	0.1
PUE	1.2	1.3	1.4
$k_i v_i$	1	1	1
β	1	1	1
$\overline{\sigma}$	1	1	1

为了验证本文方法的适用性,分别针对可再生 能源过剩与可再生能源不足两种情况设置了两种场 景,以场景1和场景2表示。

每个数据中心处于不同的电网节点,即每个数 据中心有不同的节点边际电价。目前美国 PJM 电力 市场和 ERCOT 电力市场的结算周期为 5min<sup>[36]</sup>,而 随着可再生能源的渗透,未来的结算周期可能会进 一步降低。因此,本文模拟 5min 内价格信号波动两 次。

同时,每个数据中心的所获得的可再生能源电 量也不同,由于受外界环境影响,可再生能源的出 力具有波动性,从前端传输的数据中心的计算任务 同样存在较大的波动,因此本文给可再生能源出力 以及计算任务增加随机扰动ε,以模拟可再生能源出 力及计算任务的不确定性,具体取值如表2所示:

		DC1	DC2	DC3
可再生能源(MW	)(场景1)	$20+\mathcal{E}_{1,1}$	$40 + \mathcal{E}_{1,2}$	$50 + E_{1,3}$
可再生能源(MW	)(场景2)	$30 + \varepsilon_{2,1}$	$8+\mathcal{E}_{2,2}$	$4+\mathcal{E}_{2,3}$
脑由俭故 ( 元	0s~100s	0.8	0.6	0.7
	101s~200s	0.5	0.7	0.6
/KW )	201s~300s	0.8	0.9	0.6
<b>住</b> 由	0s~100s	0.3	0.4	0.2
	101s~200s	0.4	0.2	0.2
/KW )	201s~300s	0.2	0.3	0.5
总任务量(	<b>∱/s</b> )		$3*10^{6+} \varepsilon_L$	

表2 仿真环境参数

#### 3.2 控制频率影响分析

由于可再生能源出力以及计算任务均存在较大的波动性,且难以被十分准确的预测,提升控制频率有助于降低不确定性带来的影响,本文首先对比 了控制频率对经济性的影响。基于最优控制,每次 优化可以在 100ms 之内完成求解,但考虑到调节过 程中的还存在任务分解、通信等耗时,本文取最快 的控制频率为 5s/次。

场景 1 模拟了可再生能源出力较多,总出力能 够满足数据中心能耗需求且有一定剩余的场景,场 景 2 则模拟了可再生能源出力总量不能够满足数据 中心能耗需求的情况。两种场景下 3 个数据中心任 务分配及功率情况如图 2、图 3 所示:





两种场景下,不同控制频率在 5min 中的总效益 如表 3 所示:

表 3 不同控制频率下的效益

频率	5s/次(元)	15s/次(元)	30s/次(元)	60s/次(元)
场景1	10659.16	9996.86	7254.43	6689.15
场景 2	-18897.58	-20310.85	-20002.74	-20621.53

从图 2 中可以看出,在场景 1 下,3 个数据中心 周边可再生能源的发电总量大于数据中心的能源需 求,因此除数据中心自身的能量消耗外,还可以将 多余光伏出售给电网。在 0~100s 时,数据中心 3 的 售电价格最低,因此尽可能将计算任务调度至数据 中心 3 处,以提升可再生能源的售电效益。在 101~200s 时,数据中心1处的售电价格最高,但其 可再生能源的容量较小,因此其功率依然维持在较 小的水平。在 201~300s 时,数据中心 3 处的售电价 格变为最高,从而分配给其的任务被相应减少,以 提升 3 个数据中心总的售电收益。由于外部价格有 变化,同时可再生能源出力以及计算任务有波动, 因此高控制频率能够更快的追踪不确定性带来的波 动,从而提升系统总体效益。从表3中也可以看出, 在场景1中,控制频率为5s/次时,收益为10659.16 元,相较于1min/次的控制频率,收益提升了59.35%。

图 3 所示的是场景 2 可再生能源出力不足场景 下三个数据中心的功率分配。由于可再生能源出力 较小,数据中心所需电量主要来自于外电网购电。 由于外部电网的购电价格总是高于对外售电的价 格,因此数据中心的任务分配优先使用本地的可再 生能源,不足的部分再从外电网购买。在 0~100s 时, 系统尽可能的从数据中心 2 处购电以满足功率需求, 在 101~200s 时,数据中心 2 承担了较多的计算任务, 同时由于其本身的可再生能源出力就较高,其接近 于满载运行。在 201~300s 时,则尽可能将功率分配 至购电价格较低的数据中心 3 处。由于高频率控制 的灵活性,场景 2 下的总成本也能被有效的降低。 由于可再生能源的出力具有较强的波动性且难 以被完全准确的预测,因此高频的控制能够更好的 追踪可再生能源出力的波动。本文统计了两种场景 下可再生能源就地消纳的情况,如表4所示。

表 4 不同控制频率下的可再生能源就地消纳率

频率	5s/次	15s/次	30s/次	60s/次
场景1	12.17%	4.34%	9.19%	5.93%
场景 2	97.84%	97.48%	97.91%	97.66%

从表 4 中可以看出,场景由于高频控制能够更 好的追踪可再生能源出力的波动,有效的提升了可 再生能源就地消纳率,相较于 1min/次的控制频率, 5s/次的控制频率下可再生能源的就地消纳率提升了 6.24%。而在场景 2 中,由于可再生能源出力较小, 因此不同控制频率下的就地消纳率差别不大,均在 97%以上。

场景 2 中,控制频率为 30s/次时可再生能源的就 地消纳率最高,这是由于场景 2 下各个控制频率对 可再生能源消纳的影响很小,而对于实时电价和计 算任务不确定性的追踪对总体效益的影响更大,因 此高频控制下,虽然可再生能源就地消纳率稍有降 低,但成本、计算时延等指标结果更优,从而实现 更好的总体效益。因此,本文继续就不同控制频率 下,经济性和延迟指标的优化结果进行分析。

#### 3.3 经济性与延迟优化分析

前述仿真实验验证了本文算法中计算任务调度 所带来的效益提升,在前文的理论分析中,在相同 的计算任务下,数据中心的功率与延迟存在一定的 关系,从而适当提升数据中心的计算时延,能够提 升系统总体的运营效益,本文基于场景2,又对数据 中心计算时延与系统的运行成本之间的关系进行了 仿真实验。在前述仿真实验中,延迟性能的权重取 值为1×10<sup>4</sup>,以其为基准,验证不同权重下算法控制 的效果,实验结果如图4、图5所示:





图 5 不同 γ 取值下总效益

本文算法在设计目标函数时,考虑了信息层面 服务质量这一因素,通过对目标函数中γ值的调节能 够很方便的变动经济成本和延时成本之间的权重关 系。从图 4 和图 5 中可以看出,延时的约束上限为 200ms,在γ值较小时,系统优化时主要考虑经济成 本,因此三个数据中心的计算时延普遍较高,此时 经济成本较低。而随着γ值的增加,系统在进行优化 调节时,更多的考虑了信息层面的服务质量,系统 的计算时延有了明显的下降,但随着系统的成本也 有了明显的上升。同时,对比 10 倍权重和 100 倍权 重的仿真结果可以看出,当计算时延较低时,继续 降低计算时延,需要牺牲较多的经济成本。因此, 在实际环境中,可以根据需求动态的调整权重因子 γ,以实现经济成本与时延成本之间的均衡。

# 4 结论

多个数据中心之间可以通过计算任务的分配实 现功率的转移,单个数据中心也可以通过服务器开 关、CPU 频率调节等手段协调功耗与计算时延。本 文基于最优控制理论,实现多个数据中心的任务分 配以及每个数据中心功率与计算时延的统一优化。 同时,采用动态微分方程对所求问题建模,实现数 据中心任务的快速控制。理论分析与仿真结果表明:

 1)通过多个数据中心计算任务的调度,能够有 效降低系统总体运行成本,且能够有效提升可再生 能源的就地消纳率;

2)采用动态微分方程对所求问题建模,能够实

现数据中心的秒级的优化控制,从而更好的追踪由 于可再生能源出力以及计算任务不确定性带来的波 动,以提升系统总体的经济效益;

3)在满足计算延迟约束的前提下,提高目标函数中延迟项的权重,能够有效降低计算延迟,但也 会大幅提升系统运行总成本,在实际运行中可以根据需求灵活调节权重。

综上,本文所提基于最优控制的能量信息协同 优化策略能够综合考虑能源层的能量成本以及信息 层的计算时延,实现计算任务分配及数据中心功率 控制的快速联合优化控制,降低系统运营成本,提 升可再生能源就地消纳率,并能够根据需求调整计 算时延。未来的研究可以进一步细化信息层和能量 层的研究。在能量层面,可以细化数据中心的能耗 组成,考虑冷热电等多种形式综合能源供给系统的 调节,在信息层面,可以细化计算任务,包括不同 任务之间不同的复杂程度及不同的时延需求等。

#### 致谢

本文得到了"腾讯高校合作犀牛鸟基础平台技 术专项"的支持, 谨此致谢。

# 参考文献

- [1] 李洋,肖泽青,聂松松,曹军威,华吴辰.生成对抗网络及其在新能源数 据质量中的应用研究综述[J].南方电网技术,2020,14(02):25-33.
   Li Yang,Xiao Zeqing,Nie Songsong,Cao Junwei,Hua Haochen. Review of Research on Generative Adversarial Network and Its Application in New Energy Data Quality [J]. Southern Power System Technology,2020,14(02):25-33. (in Chinese)
- [2] H. Hua, C. Hao, and Y. Qin. Internet thinking for layered energy infrastructure[M]. A. Zobaa, J. Cao (eds). Energy Internet: Systems and Applications. Switzerland AG: Springer Nature, 2020: 421-437.
- [3] H. Hua, Y. Qin, C. Hao, and J. Cao. Stochastic optimal control for energy Internet: A bottom-up energy management approach[J]. IEEE Transactions on Industrial Informatics, 2019,5(3): 1788-1797.
- [4] J. Cao, H. Hua, and G. Ren. Energy use and the Internet[M]. The SAGE Encyclopedia of the Internet. Newbury Park, CA, USA: Sage, 2018: 344-350.
- [5] 绿色和平组织,华北电力大学.点亮绿色云端:中国数据中心能耗与 可再生能源使用潜力研究[R].北京:绿色和平组织,华北电力大学, 2019.
- [6] 王继业,周碧玉,刘万涛,虎嵩林.数据中心跨层能效优化研究进展和发展趋势[J].中国科学:信息科学,2020,50(01):1-24.
   Wang Jiye, Zhou Biyu, Liu Wantao, Hu Songlin. Research progress and

development trend of cross-layer energy efficiency optimization in data centers[J]. Scientia Sinica(Informationis), 2020, 50(01):1-24. (in Chinese)

- [7] 冯成,王毅,陈启鑫,张轩,罗钢.能源互联网下的数据中心能量管理综述[J].电力自动化设备,2020,40(07):1-9.
   Feng Cheng,Wang Yi,Chen Qixin,Zhang Xuan,Luo Gang. Review of energy management for data centers in energy internet[J]. Electric Power Automation Equipment, 2020, 40(07): 1-9. (in Chinese)
- [8] Yao F, Demers A, Shenker S. A scheduling model for reduced CPU energy[C]. Proc of the 36th Annual Symp on Foundations of Computer Science. Piscataway, NJ: IEEE,1995:374-382.
- [9] Albers S. Energy-efficient algorithms[J]. Communications of the ACM, 2010,53(5):86-96.
- [10] 李文信,齐恒,徐仁海,周晓波,李克秋.数据中心网络流量调度的研究 进展与趋势[J].计算机学报,2020,43(04):600-617.
  Li Wenxin, Qi Heng, Xu Renhai, Zhou Xiaobo, Li Keqiu. Data Center Network Flow Scheduling Progress and Trends[J]. Chinese Journal of Computers, 2020, 43(04): 600-617. (in Chinese)
- [11] 吴刚,高赐威,陈宋宋,李德智,刘强.考虑需求响应的数据中心用电负 荷优化研究综述[J].电网技术,2018,42(11):3782-3788.
  Wu Gang,Gao Ciwei,Chen Songsong,Li Dezhi,Liu Qiang. A Survey on Data Center Power Load Optimization Considering Demand Response[J]. Power System Technology,2018,42(11):3782-3788. (in Chinese)
- [12] 高赐威,曹晓峻,闫华光,苏卫华.数据中心电能管理及参与需求侧资源 调度的展望[J].电力系统自动化,2017,41(23):1-7.
  Gao Ciwei, Cao Xiaojun, Yan Huaguang, Su Weihua. Energy Management of Data Center and Prospect for Participation on Demand Side Resource Scheduling[J]. Automation of Electric Power Systems,2017,41(23):1-7. (in Chinese)
- [13] 宋杰,孙宗哲,刘慧,鲍玉斌,于戈.混合供电数据中心能耗优化研究进展[J].计算机学报,2018,41(12):2670-2688.
   Song Jie,Sun Zongzhe,Liu Hui,Bao Yubin,Yu Ge. Research Advance on Energy Consumption Optimization of Hyper-Powered Data Center[J]. Chinese Journal of Computers,2018,41(12):2670-2688. (in Chinese)
- [14] L. Yu, T. Jiang and Y. Zou. Real-Time Energy Management for Cloud Data Centers in Smart Microgrids[J]. IEEE Access, 2016,4: 941-950.
- [15] L. Yu, T. Jiang and Y. Zou. Distributed Real-Time Energy Management in Data Center Microgrids[J]. IEEE Transactions on Smart Grid, 2018,9(4): 3748-3762.
- [16] 曹晓峻,高赐威,李德智,杨建林.数据网络与电力网络混合运行建模及 其参与系统经济运行[J].中国电机工程学报,2018,38(05):1448-1456. Cao Xiaojun, Gao Ciwei, Li Dezhi, Yang Jianlin. Mixed Operation Model of Data Network and Power Network and Its Participation in the Economic Operation of Power System [J]. Proceedings of the CSEE, 2018,38(05):1448-1456. (in Chinese)
- [17] 陈敏,高赐威,陈宋宋,李德智,刘强.考虑数据中心用电负荷调节潜力 的双层经济调度模型[J].中国电机工程学报,2019,39(05):1301-1314. Chen Min, Gao Ciwei, Chen Songsong, Li Dezhi, Liu Qiang. Bi-level Economic Dispatch Modeling Considering the Load Regulation Potential of Internet Data Centers [J]. Proceedings of the CSEE, 2019,

39(05): 1301-1314. (in Chinese)

- [18] 王晴,刘友波,黄杨,刘畅,陈刚,刘俊勇.考虑数据中心需求响应的城市 电网阻塞管理[J].电网技术,2020,44(08):3129-3138.
  Wang Qing, Liu Youbo, Huang Yang, Liu Chang, Chen Gang, Liu Junyong. Congestion management of Urban Power Grid Considering Demand Response of data center [J].Power System Technology, 2020,44(08):3129-3138. (in Chinese)
- [19] 高赐威,吴刚,陈宋宋.考虑地理分散的数据中心服务器频率调节的电网降损模型[J].中国电机工程学报,2019,39(06):1673-1681+1863.
  Gao Ciwei,Wu Gang,Chen Songsong. A Model Aimed at Reducing Power Net Loss Considering Frequency Scaling of Servers in Geo-Distributed Data Centers[J]. Proceedings of the CSEE, 2019, 39(06):1673-1681+1863. (in Chinese)
- [20] L. Yu, T. Jiang, and Y. Zou, "Distributed online energy management for data centers and electric vehicles in smart grid," IEEE Internet Things J., vol. 3, no. 6, pp. 1373–1384, 2016.
- [21] 赵剑,黄悦琛,李海阳,何湘粤.垂直起降运载火箭返回轨迹不确定性优化[J].航空学报,2021,42(11):262-276.
  Zhao Jian, Huang yuechen, Li Haiyang, he Xiangyue. Uncertainty optimization of return trajectory of vertical takeoff and landing launch vehicle [J] Acta Aeronautica et Astronautica Sinica, 2021,42 (11): 262-276. (in Chinese)
- [22] 黄少锋,李慧,李轶凡,张月品.换相序技术应用于稳定控制的条件及最优控制策略[J].电工技术学报,2021,36(11):2245-2254.
  Huang Shaofeng, Li Hui, Li Yifan, Zhang yuepin Conditions and optimal control strategy of commutation sequence technology applied to stability control [J] Transactions of China Electrotechnical Society, 2021,36 (11): 2245-2254. (in Chinese)
- [23] 李奇,王晓锋,孟翔,张国瑞,陈维荣.基于在线辨识和极小值原理的 PEMFC 混合动力系统综合能量管理方法[J].中国电机工程学 报,2020,40(21):6991-7002.

Li Qi, Wang Xiaofeng, Meng Xiang, Zhang Guorui, Chen Weirong Integrated energy management method of PEMFC hybrid system based on on-line identification and minimum principle [J] Proceedings of the CSEE, 2020,40 (21): 6991-7002. (in Chinese)

- [24] Google, "Cluster workload traces," 2015(2018). [Online]. Available: https://github.com/google/cluster-data
- [25] Peter Asmus. Data Centers and Advanced Microgrids[R]. Boulder: Navigant Consulting, 2017.
- [26] X. Lu, F. Kong, X. Liu, J. Yin, Q. Xiang and H. Yu. Bulk Savings for Bulk Transfers: Minimizing the Energy-Cost for Geo-Distributed Data Centers[J]. IEEE Transactions on Cloud Computing,2020,8(1): 73-85.
- [27] Howard Cheung, Shengwei Wang, Chaoqun Zhuang, Jiefan Gu. A simplified power consumption model of information technology (IT)

equipment in data centers for energy system real-time dynamic simulation[J]. Applied Energy,2018, 222: 329-342.

- [28] M. Dayarathna, Y. Wen and R. Fan. Data Center Energy Consumption Modeling: A Survey[J]. IEEE Communications Surveys & Tutorials,2016,118(1): 732-794.
- [29] A. Khosravi, L. L. H. Andrew and R. Buyya, Dynamic VM Placement Method for Minimizing Energy and Carbon Cost in Geographically Distributed Cloud Data Centers[J]. IEEE Transactions on Sustainable Computing, 2017,2(2): 183-196.
- [30] Liu Z, Lin M, Wierman A, et al. Greening geographical load balancing[J]. IEEE/ACM Transactions on Networking, 2015, 23(2): 657-671.
- [31] L. Chen, S. Zhou and J. Xu. Computation Peer Offloading for Energy-Constrained Mobile Edge Computing in Small-Cell Networks[J]. IEEE/ACM Transactions on Networking, 2018, 26(4): 1619-1632.
- [32] Ahmad F, Vijaykumar T N. Joint optimization of idle and cooling power in data centers while maintaining response time[C]. Fifteenth Edition of ASPLOS on Architectural Support for Programming Languages and Operating Systems. ACM, 2010: 243-256.
- [33] N. Liu, X. Yu, C. Wang, C. Li, L. Ma and J. Lei. Energy-Sharing Model With Price-Based Demand Response for Microgrids of Peer-to-Peer Prosumers[J]. IEEE Transactions on Power Systems,2017,32(5): 3569-3583.
- [34] Frank L. Lewis, Draguna L. Vrabie, and Vassilis L. Syrmos. Optimal control[M]. New Jersey: John Wiley & Sons, 2012.
- [35] Beal, L.D.R., Hill, D., Martin, R.A., and Hedengren, J. D., GEKKO Optimization Suite [J]. Processes, 2018,6(8):106.
- [36] 武昭原,周明,王剑晓,汤文渊,元博,李庚银.双碳目标下提升电力系统 灵活性的市场机制综述 [J/OL].中国电机工程学 报:1-18[2022-01-14].DOI:10.13334/j.0258-8013.pcsee.212117.
  Wu Zhaoyuan, Zhou Ming, Wang Jianxiao, Tang Wenyuan, Yuan Bo, Li Gengyin Overview of market mechanism for improving po wer system flexibility under double carbon target [J / OL] Proceed ings of the CSEE: 1-18 [2022-01-14] DOI:10.13334/j.0258-8013.pc see.212117. (in Chinese)