能源互联网下的数据中心能量管理综述

冯 成[1],王 毅[2],陈启鑫[1],张 轩[3],罗 钢[3]
(1. 清华大学 电力系统及发电设备控制和仿真国家重点实验室,北京 100084;
2. 苏黎世联邦理工学院,瑞士 苏黎世 8092;
3. 广东电网有限责任公司 电力调度控制中心,广东 广州 510699)

摘要:数据中心在大数据时代扮演重要角色,随着数据中心的持续铺开建设,其高能耗问题日益凸显,能量管理由此成为了数据中心产业关注的焦点。在能源互联网环境下,数据中心将会拥有更多样化的供能选择,可以发掘更灵活高效的能量管理方式,其仅依靠电能的能量管理模式将发生根本性变革。从能源互联网视角出发,对数据中心能量管理相关研究和实践进行了综述。首先梳理了数据中心的结构层次与设备组成,总结了数据中心的能耗结构、能量效率和能耗模型;然后在此基础上深入分析了数据中心在能源互联网背景下的能量管理方法,并总结了国内外相关的数据中心项目实践;最后探讨了未来实现能源互联网背景下数据中心能量管理的挑战和难点。

关键词:数据中心;能源互联网;能量管理;能耗模型;需求响应
中图分类号:TM 73 **文献标志码**:A DOI:10.16081/j.epae.202007003

0 引言

建设能源互联网是推动我国能源革命的重要途径之一,能源互联网能够最大限度地提高能源利用效率,引导能源消费向着集约化、清洁化的方向发展[1-2]。除了能源设备的互联,能源互联网也着眼于海量分布式设备之间的信息交互与协调,以实现能量与信息的双向流动[3]。海量信息的存储、计算需要依赖于强有力的信息设备的支持,而信息设备中发展最迅速的是具有互联网规模范围的数据中心[4]。随着数据中心的不断铺开建设,其能耗已经上涨到十分可观的水平。2015年,国家能源局《关于印发国家绿色数据中心试点工作方案的通知》指出,我国数据中心发展迅猛,总量已经超过40万个,数据中心年耗电量超过全社会用电量的1.5%[5]。据预测,到2020年,数据中心预计将消耗约8%的全球电力。而对于数据中心运营商而言,能耗费用也是其主要的运营支出,单个数据中心运营商每年所支付的能耗费用可能高达上千万元[6]。在高额的能耗费用面前,数据中心的发展裹足不前,因此实现高效的能量管理、减少能耗费用成为了数据中心产业亟需迫切解决的问题。

数据中心消耗电能属于电力负荷,同时配备有本地备用电源可以自产电能,另外,数据中心工作时产出热能并需要冷源给予支撑,并且未来有望引入燃气作为冷热电三联供CCHP(Combined Cooling Heating and Power)、燃料电池的一次能量供应来源,所以数据中心并非单一电能的消费者,而是多种形式能源的产消者。随着能源互联网的纵深发展,多能协同给数据中心的能耗管理带来了新机遇,其原有的仅依靠电能的供能模式会发生根本性变革。那么能源互联网究竟会催生出哪些新的数据中心的能源供应选择?数据中心运营商应该如何把握新选择带来的机遇,创新能量管理方式以节省能耗成本?应用这些能量管理方法又会面临哪些挑战?

为回答上述问题,本文从能源互联网的视角出发,对数据中心能量管理的相关研究和实践进行了综述。首先梳理了数据中心的结构层次与设备组成,总结了数据中心的能耗结构、能量效率和能耗模型;在此基础上深入分析了数据中心在能源互联网背景下的能量管理方法,并总结了国内外相关数据中心项目实践;最后探讨了未来实现能源互联网背景下数据中心能耗管理的挑战和难点。

收稿日期:2020-03-10;**修回日期**:2020-05-11
基金项目:国家重点研发计划项目(2016YFB0900100);广东省重点领域研发计划项目(2019B111109002);广东电网科技项目(036000KK52180031)
Project supported by the National Key R&D Program of China(2016YFB0900100),Guangdong Province R&D Projects in Key Areas(2019B111109002) and Guangdong Power Grid Science and Technology Project(036000KK52180031)

1 数据中心简介

1.1 服务架构层次

数据中心指用于安置大规模计算机/服务器系统及相关部件的设施。通常而言,一个数据中心服务商会拥有多个地理位置分布各异的数据中心实体,每个实体又由多台服务器组成,多个数据中心通过网络连接形成数据中心网络[7]。数据中心网络的基础架构如图1所示。其一般为三层树形拓扑架构,3个层次分别为核心层(网络的高速交换主干)、

汇聚层(提供基于策略的连接)、接入层(将数据中心设备接入网络),每一层网络通过不同类型的路由器和通信线路相互连接[8],不同地理分布的数据中心实体可以通过可靠的通信链路如光纤等进行信息传送,数据中心之间的信息数据交换是其最基本的功能之一。接入层的服务器分为前端和后端,前端服务器旨在处理所有来自数据中心用户的通信,数据中心用户通过前端服务器访问各数据中心,而后端服务器进行存储计算服务。

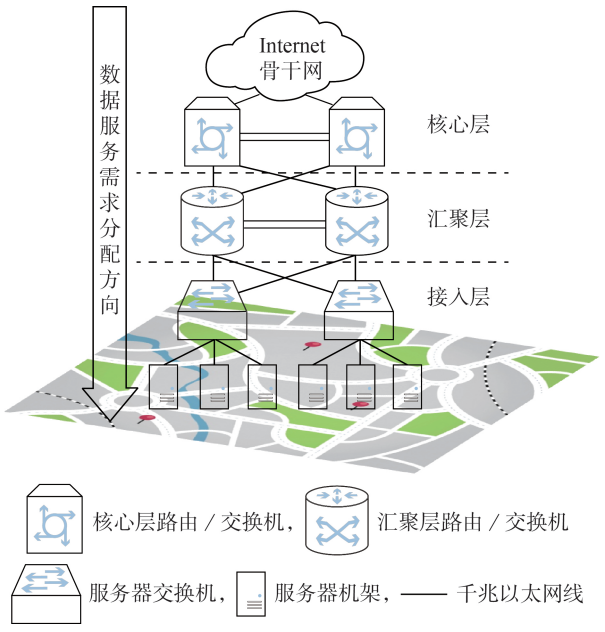

图 1 数据中心服务架构层次

Fig.1 Hierarchy of service structure for data centers

在实际工作中,当有数据服务需求请求通过核心骨干网到达前端服务器时,前端服务器将数据服务需求分配到地理分布各异的数据中心实体执行。一个数据中心服务商拥有的数据中心实体的个数为数十个到数百个,而一个数据中心所管理的服务器的数量级在数百到数万之间,总体而言,数据中心的服务器总个数可以达到千甚至百万级别[6,9]。每个数据中心实体和某个能源网区域相关联,一般认为如果数据中心是地理分散的,则其分散程度较高,每个数据中心的能价各异,甚至每个数据中心所在区域的电网运营商也会不同[6,10-11]。

数据中心的运营模式一般分为自管式和租赁式这2种[11]。自管式数据中心服务商拥有对所有数据中心设备的直接控制权,故其可以自主决策所有服务器的能耗调节策略;而租赁式数据中心服务商将不同的数据中心托管给不同的用户,在租赁后数据服务商难以对数据中心进行直接操作,在参与需求响应时,数据中心服务商需要激励用户进行能耗调节以完成总的能量管理目标而不能单方决断[12-13]。

1.2 硬件设备组成

数据中心硬件设备组成如表1所示。单个数据中心实体的内部一般由电力设备、制冷设备、信息设备组成。电力设备主要包括电源转换装置(AC-DC转换器和DC-AC转换器)、不间断电源(UPS),随着新能源技术和储能技术的发展,也有一些数据中心开始建设多种备用能源(现有的大多为柴油发电机)和专用的储能电池装置;制冷设备主要包括用于冷却数据中心机房的空调、各种冷热泵和冷循环设备,其中通过水循环进行冷却是现有的较为成熟的制冷方案;信息设备主要包括数据中心的服务器、存储器、通信设备等[14],其分别负责数据任务的计算、存储和不同节点间的通信。

表 1 数据中心实体的硬件设备组成

Table 1 Components of hardware infrastructures of data centers

设备类型	设备名称
电力设备	电源转换装置、不间断电源、本地备用电源、本地电储能
制冷设备	机房空调系统、冷热水 / 气泵、冷循环设备、储冷储热罐
信息设备	服务器、存储器、通信设备

2 数据中心的能耗

2.1 能耗组成与能耗效率

对于数据中心具体的能耗组成,电力设备(仅电能转换设备)、制冷设备、信息设备的耗电量占比分别约为10%、40%、50%。其中在信息设备的耗电组成中,服务器的耗电量最高,大约占整个数据中心耗电量的40%,其余信息设备如存储器和通信设备的耗电量占10%。单个服务器机柜的功率在2~10 kW之间,考虑到一个数据中心的服务器鲜有同时全部满载工作的情况,一个数据中心服务商所拥有的设备耗电量一般在MW级[15-16]。

电能利用效率PUE(Power Usage Effectiveness)是衡量数据中心能源效率的重要评价指标,其计算公式为:

$$\lambda_{PUE}=E_t/E_{IT} \tag{1}$$

其中,λ_{PUE}为数据中心的PUE值;E_t为数据中心全年总耗电量;E_{IT}为数据中心信息设备的全年总耗电量。国务院在《国务院关于促进云计算创新发展培育信息产业新业态的意见》中以及工信部在[2013]48号通知中均将绿色数据中心定位为λ_{PUE}在1.5以下[17]。而国内数据中心的平均$\lambda_{PUE}=2.2$,Google公司于2016年宣布其数据中心的λ_{PUE}可达1.08,可见国内数据中心与国际先进数据中心的水平有较大的差距[18]。现有的数据中心机房设计方案中大多强调需要对机房的碳排放和能量消耗情况进行监控并对PUE进行管理,这已经为数据中心进行能量管理提供了良好的硬件基础。

PUE可用于估计数据中心的总能耗。以国内典型数据中心服务商为例,其拥有3 000个服务器机柜,单个机柜的装机功率为4.4 kW,国内PUE的平均值为2.2,按平均负荷率为50% 进行计算,该数据中心服务商所拥有的设备总平均功率为14.52 MW。电费按0.7元/(kW·h)(国内较低水平的工业电价)进行计算,其日电费成本为24.4万元,年电费成本高达8 906万元(一年按365天进行计算)。如此高的运营成本使得数据中心的成本回收周期十分漫长甚至会有入不敷出的情况发生[9]。

值得注意的是,传统的PUE将电能消耗量作为指标对数据中心进行评价。然而在能源互联的背景下,其能量来源不仅仅是电网,传统的PUE不再适用,需要考虑多能源的PUE。一种做法是考虑能源转换系数的PUE,具体如下:

$$\lambda_{\mathrm{PUE}}=(E_{\mathrm{ee}}+E_{\mathrm{oe}}\eta)/E_{\mathrm{IT}} \tag{2}$$

其中,E_{ee}为数据中心消耗的电能;E_{oe}为数据中心消耗的其他能源的能量;η为能源转换系数。

2.2　能耗模型

2.2.1　电负荷能耗模型

电负荷模型主要指服务器的能耗模型[19-20]。较为典型的是采用动态电压频率技术的CPU能耗调节模型,其基于服务器运行时的内在物理机理,即运行时CPU级能耗与工作频率、工作电压、CPU利用率之间定量关系来计算能耗,具体如式(3)所示。

$$P_{\mathrm{CPU}}(U)=P_{\mathrm{CPU,min}}+CfV^2U \tag{3}$$

其中,$P_{\mathrm{CPU,min}}$为CPU空闲时的最小能耗;f为CPU电压频率;V为电压幅值;U为CPU使用率;C为常数,其由CPU的硬件特性决定,一般通过实验测得。

在大量服务器集群的情况下,难以建立和电量直接相关的能耗模型,此时宜采用统计拟合的方式,即假设多台服务器的总能耗是服务器利用率的函数,在获得运行数据后再对具体关系进行拟合。一般为了方便处理,假设关系函数是线性函数,即服务器可以处于空闲到满载的任意状态,随着数据服务需求大小的变化呈理想线性变化,具体如式(4)所示。

$$P_{\mathrm{serv}}(u)=P_{\mathrm{serv,min}}+u(P_{\mathrm{serv,max}}-P_{\mathrm{serv,min}}) \tag{4}$$

其中,$P_{\mathrm{serv,min}}$为无数据负载时的服务器能耗;$P_{\mathrm{serv,max}}$为满载时的服务器能耗;u为服务器的利用率,由当前的数据负载量除以最大负载量得到。

以上2种解析化的方法因为易处理而被广泛采用,然而确定相关参数需通过服务器的长期运行数据拟合得到,获取周期长且难度较高。此外,由于模型可能过于理想化,与实际运行工况偏差可能较大。

2.2.2　热负荷能耗模型

与电负荷类似,热负荷的建模思路可分为基于内在物理规律的建模方式和基于统计的建模方式。其中基于内在物理规律建立的模型又可分为基于流体理论的机房风扇模型和基于热传导理论的冷却空调模型。

机房风扇模型以风扇功率与风扇转数的立方的正比关系($P_{\mathrm{fan}}\propto s_{\mathrm{fan}}^3$)[21]规律为基础,且当服务器功率越大时,风扇转速也需要相应正比增大以增加散热量($P_{\mathrm{serv}}\propto s_{\mathrm{fan}}$),故有如式(5)所示的热功率关系。

$$e_{\mathrm{fans}}(P_{\mathrm{serv}})=kP_{\mathrm{serv}}^3 \tag{5}$$

其中,e_{fans}为风扇散热所需要的驱动功率,也可以看作无散热时的产热量;k为与温度相关的参数,其由室内温度和室外温度的差值决定,可以通过热传导理论进行计算。

冷却空调模型考虑了散热过程中的热流动态过程,较为复杂的有热阻等效模型。相似模型中大部分具有如式(6)所示的形式。

$$e_{\mathrm{crac}}=\phi_{\mathrm{crac}}\Delta T \tag{6}$$

其中,e_{crac}为冷却空调所需的总能耗;ϕ_{crac}为用于冷循环的热流体密度、比热容、流量的乘积;ΔT为装置流出的冷气/水流与返回装置的热气/水流之间的温差。式(6)体现了当热流体密度、比热容、流量变化不大时,热负荷可以看作降温温差的一次函数[22]。

基于统计的建模方式利用数据直接拟合数据中心所需冷却功率与服务器利用率之间的关系,典型地是以分段非线性函数的形式出现:

$$e_{\mathrm{cool}}=\begin{cases}\alpha u^2+\beta u+\gamma+\theta & u<25\%\\ \theta & \text{其他}\end{cases} \tag{7}$$

其中,e_{cool}为冷却设备的总能耗;α、β、γ、θ为拟合的参数值。上述模型假设当服务器利用率低于25% 时,整个热负荷是可以调整的,而当其大于等于25% 时,为保证服务器中半导体设备的正常工作,热负荷是不可调的。

3　数据中心的能量管理方法

在能源互联网下,数据中心的设备架构如图2所示,其用能、供能向着能源形式多样化的方向发展,并且其可以参与到广域能源互联的互动中。本节将概述能源互联网下数据中心的能量管理方法,具体的方法分类如表2所示。本节讨论的管理方法指的是数据中心建成之后所能采取的措施,一些在数据中心建成后就已尘埃落定的调节方式(如电源特性、电能转换特性、芯片能效、制冷设备能效)不在考虑范围内。

3.1　本地设备级能量管理

本地设备级能量管理是指直接对设备的工作状态进行操作进而降低能耗的管理方式。

3.1.1　服务器功率调整

服务器运行时处于休眠状态与满负荷状态之间的状态,其运行时的能耗由式(3)和式(4)给出。运行时,由于休眠状态下不进行运算也需要消耗电能,

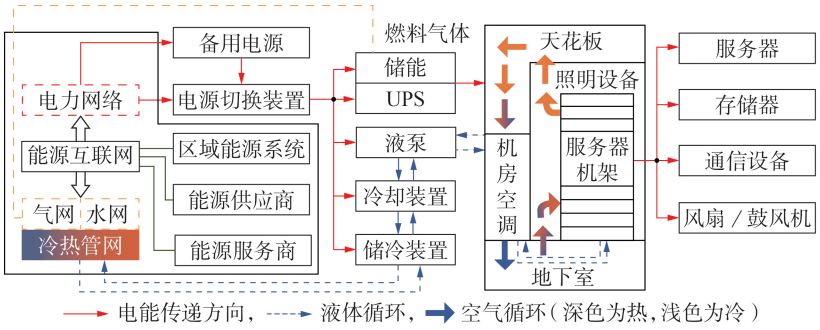

图2 能源互联网下的数据中心架构

Fig.2 Architecture of data centers in energy internet

表2 能量管理方法分类

Table 2 Categories of energy management methods

方法类别	管理方法
本地设备级	服务器功率调整 本地取能与储能 能量转换与回收
数据服务级	数据服务空间分配 数据服务时间转移 策略降低服务质量
能源网络级	参与能源需求响应 利用广域多能互补

应该尽量关停不必要的服务器，或者改变CPU电压大小和频率以增加CPU的使用率[23-28]。

服务器功率调整是最直接的调整数据中心能耗的方式，然而实施起来却比较困难。一方面是由于数据负荷具有随机性而数据中心的启停同样需要时间，难以实时动态调整跟踪数据服务需求；另一方面是调整实际运行节点数或者改变CPU运行条件均会影响到整个系统的运行效率和稳定性，对于数据服务而言存在较大的风险。

3.1.2 本地取能与储能

数据中心作为能耗大户，其被要求减少运营时的碳排放量，因此越来越多的数据服务商逐渐将可再生能源电源作为其备用供电电源，其中比较成熟的方案是利用环境中的风、光资源进行发电，同时配备有一定容量的储能来应对风光资源的间歇性问题[28-31]。以上方法实际上使得数据中心成为了一个数据中心微电网。

利用本地冷源是数据中心冷能管理最直截了当的方式。现有的数据中心冷源大多来源于自来水网的冷水，或自建冷却塔和冷却机组进行冷却。有研究利用自然冷源设计数据中心的板热换热器，也可因地制宜利用自然湖水冷源进行冷能循环供应。另外，若有储冷设备，可以提前管理设备蓄冷过程（包括水蓄冷和冰蓄冷）[26,32]，以备冷能供给紧张时使用。由于某些数据中心要求不间断供冷，实际上蓄冷会占用供冷系统的一部分储冷应急容量，然而占用此部分应急容量用于能量管理会对系统的可靠性造成影响，所以在实际应用中需要充分权衡。

3.1.3 能量转换与回收

能量转换与回收主要指在数据中心本地回收利用数据中心产生的余热，最典型的是在数据中心本地建设冷热电三联供系统。冷热电三联供系统通过燃气发电机冷热电联供，利用燃料的化学能生成烟气的热能，按品质分别转化为电能和热能，进而实现对燃气的高效梯级利用[33]。一般情况下，数据中心本地并没有充足的燃气资源，冷热电三联供系统需要从能源互联网中购买燃气，这要求数据中心运营商在多种功能方案中进行成本和效益的对比，当燃气供能的综合效益是最佳时适宜采用这种能量管理方式[34]。从2015年开始，由于非居民用天然气连续降价，利用天然气对数据中供能的效益明显提高，国内大型数据中心开始逐步采用燃气冷热电三联供系统[35]。

3.2 数据服务级能量管理

数据服务级能量管理是指对数据服务需求进行操作而降低能耗的管理方式。

3.2.1 数据服务空间分配

对于数据服务需求，其一般被分为批处理数据服务需求和交互型数据服务需求。前者对完成时间的要求较低，只需要在规定的时间内完成即可，典型的如索引、数据挖掘；而后者对响应的实时性要求较高。转移、迁移、分配数据服务需求一般都是指批处理服务需求。

数据服务需求空间分配又被称为数据服务需求的地理负载平衡，由于数据中心中拥有多个地理分布各异的数据中心，每个数据中心的能价可能存在差异，将批处理数据服务需求分配到能价较低的数据中心，利用空间上的数据负荷转移也能够降低数据中心的运营成本[6,12,26,36-45]。具体的数据服务需求分配可以以任务为最小单位，也可以以服务器上运行的虚拟机为最小单位。

3.2.2 数据服务时间转移

数据服务需求空间分配是利用地理位置上能价的不同来降低数据中心服务商的运营成本，而数据

服务需求时间转移是利用时间上能价的不同来降低运营成本,其核心思想是将数据服务需求尽量延迟到低能价的时段[24,29,39,43,46-47]。数据服务需求时间转移受到数据服务需求响应时间的限制,相较于空间分配更大。

若空间上和时间上同时存在能价差异,综合空间分配和时间延迟,可以实现数据服务需求的跨时间–空间能量管理。更进一步,在储能系统、储冷系统的支持下,不仅只有转移数据服务需求可使数据中心服务商获利,用电储能和储冷也可实现不同时段能价的套利,实现非信息设备耗能的时间转移。

由于时间转移是建立在能价存在差值的基础上,需要明确知道能价的高低才能进行管理,故能价的预测是必不可少的,而预测又会存在误差,这使得数据中心服务商面临能价波动的风险。

3.2.3　策略降低服务质量

数据服务需求在执行时有服务质量指标或者计算结果精确度的指标。服务质量规定了数据服务需求允许的延迟等,然而服务质量很大程度上是人的主观感受,并不是一个严格的执行标准,为了节省耗能费用可以策略性地降低服务质量,放宽实际的时延标准,从而在时间、空间上数据服务需求能获得更大的转移空间;而对于可以接受近似结果的数据服务需求,如搜索或者仿真,执行部分数据服务需求可以获得的精确度虽然较完全执行低,但如果能够满足一些需求,也可以纳入策略降低服务质量的范畴[30,48-50]。

策略降低服务质量针对的数据服务需求的类型较少,另外数据中心服务商需要考虑到在降低服务质量后,虽然通过能量管理可以降低运营成本,但可能也会导致数据用户不满意而退出数据服务,降低其运营收入,故需要仔细对两者做出权衡。

3.3　能源网络级能量管理

能源网络级能量管理是指对数据中心接入广域能源互联网后与网络中其余主体进行互动的管理。

3.3.1　参与能源需求响应

数据中心接入能源互联网后,可以参加能源网的需求响应项目以减少其能耗费用。此类研究现主要集中在参与电网的需求响应,对热的讨论较少,但可以进行类比拓展。在这些研究中,既包含价格型的需求响应,也有激励型的需求响应。文献[25,28,30-31,36,46,51]考虑了数据中心服务商利润和成本的权衡问题,包括电费支出、数据服务盈利[30]、储能套利[31]、本地自发电支出[51]、碳排放费用[36]和服务器容量固定投入[46];文献[38,41-43,45,52-53]讨论了数据中心所面临的电价、数据服务需求不确定性,利用随机管理[45]、两阶段管理[43]等方式对相应策略进行调整;文献[27]探讨了数据中心参与辅助服务的框架;除此之外,文献[13,54]还探讨了一些定制合同型的激励设计。

在设计需求响应机制时,需要着重考虑数据中心服务商和能源运营商之间的互动,并在能够定量衡量这种影响的情况下设计最优的激励使得电网运营者和数据中心服务商均能获利,形成双赢。

3.3.2　利用广域多能互补

数据中心工作时其数据设备会产生大量的热量,可以将数据中心接入区域供热系统/建筑供热系统的暖通空调系统,用于区域/建筑供热,实现余热回收利用[21,35]。当从能源互联网中获取不同能源存在价差时,数据中心运营商可选择最合算的能源供应形式,并可用不同能源价格进行套利[55]。如天然气价格相较于电价而言能耗代价更低,效益更优,利用本地燃气发电机组即可发电,并且可充分利用数据设备产生的余热而不减少购电以节省能耗成本。

另外,广泛的能源互联归于数据中心更多样化的冷源获取途径,如来自于液化天然气管网的冷源,将数据中心接入液化天然气管网或者利用相变材料运输液化天然气冷能,利用乙二醇提取液化天然气冷能用于数据中心冷却[56-57]。

除了冷、热负荷之外,数据中心的储能也具备多种能源选择,燃料电池取代电化学电池是数据中心储能发展的一大趋势,燃料电池相较于一般的储能有更长的寿命,其中比较有前景的是用氢燃烧和水电解的方式以实现数据中心的储能充放电。据估计,一个 20 MW 的数据中心每天大约需要 30 t 氢气[58]。由于需要补充燃料电池中的氢气,数据中心可接入现有氢运输管道,也可用运输车进行运输。

4　数据中心的能量管理实践

4.1　国内相关实践

我国地理疆域广阔,众多数据中心运营商将数据中心选址在高海拔的低温地带,其中以贵州为代表,其夏季温度低,湖泊众多,有众多的自然冷源,且水电资源丰富,电价较低。移动通信商如中国移动、通信设备巨头如华为均将数据中心落户贵州[59]。除此之外,阿里巴巴公司将数据中心选址在千岛湖旁,利用自然冷源提供冷能[60]。

2016年以来,国家发布相关政策鼓励燃气在数据中心的应用[61]。这些数据中心能源站基本上均采用内燃机加烟气热水型溴化锂机组,电力采用自发自用孤网运行模式,为了保证数据中心的可靠供电,还有一路市电作为备用。目前建成的冷热电三联供系统的市电备用采用的是“并网不上网”的模式,只能购电不能售电,不过未来这种限制有望放宽[62]。典型的如上海腾讯青浦数据中心,其天然气供电量占整个数据中心供电量的39%,制冷量占30%[63]。除此之外,由于广州市数据中心需求量大,然而其地

处低纬度地区，尤其是夏季天气炎热，对冷能供应要求较为迫切，所以广州市政府积极推动冷热电三联供项目，且广州具备利用液化天然气冷能的地理优势，如广州超算一期工程实现了冷电联供，能源综合利用效率超过70%[64]。

4.2 国外相关实践

国外数据中心运营商尤其是美国的科技巨头在利用自然冷源进行数据中心供冷走在世界前列，谷歌公司利用自然海水进行冷却，微软则直接建造了水下数据库[18]。这些利用自然冷却的数据中心都具备可观的PUE，保有国际上低PUE的记录。高纬度区域国家如芬兰和丹麦，由于全年温度较低，实现了数据中心的自然空气冷却[57]。

国外对于数据中心的余热回收利用也展开了广泛实践，特别是将数据中心的产热通过热泵/换热器接入区域能源互联网进行家庭、办公室、恒温泳池的供热。这种做法也主要应用在全年大部分时间需要供暖的高纬度地区，如Facebook在丹麦建造的数据中心，其由风电供能，利用自然海水冷却，每年还可回收2 800 MW·h的热能用于恒温泳池的供热[65]。

随着燃料电池技术的快速发展，国外已有数据中心尝试综合热电联产技术与燃料电池技术进行能耗管理的实例。有相关项目在以天然气驱动热电联产的基础上，辅有固体氧化物燃料电池支撑热电联产。如eBay在美国进行了相关实践，其所建的综合项目预计每年将提供1.75×10^6 kW·h的电能给本地数据中心[66]。

5 数据中心能量管理的挑战与研究难点

5.1 服务并发与地理分散特性

数据服务在时间上并行执行，在空间上分散分布。每个数据中心实体地处不同的自然环境，可能和不同的能源服务商、运营商互动，且面对不同的能价，这对数据中心能量管理提出了高并行、可分布以及可扩展的要求。在这个背景下，分布式算法成为了数据中心能量管理方法设计的首选，以实现能量管理的分散并行决策。在分散决策的基础上，也需要进行不同节点的协调，以达到整个数据中心尺度上的最优，即尽量靠近全局能量管理决策最优解。

现有的分布式算法依赖于将一个集中式的算法进行分解拆分，在不同节点之间进行通信和计算迭代，最后得出结果，这种模式会给计算节点通信、信息同步带来较大的开销。能够尽量减少通信开销的分布式算法是未来数据中心需求响应分布式算法的趋势。

5.2 数据动态服务质量的限制

数据服务具有各异的服务质量要求如时延、抖动等，并且有强随机性。能量管理需要实时集合随机到达服务器的数据服务需求来进行能量管理，且需要接受实时的能价信号再进行响应，响应需要考虑到此时此刻数据中心数据服务的部署状态，而此状态是实时动态变化的，以上动态、实时的调整需求都依赖于在线调整。

另外，由于数据中心的能量管理问题本身也是一种计算任务，其结果决定了用户数据服务需求的分配方式，此问题需要尽快解决，否则可能会造成数据服务处理队列的拥塞，以至于本末倒置，导致能量管理问题挤占系统过多的计算资源和通信资源。

5.3 联网多主体的互动与博弈

数据服务质量特性限制了能量管理问题的求解速度，能量管理问题需要快速在线求解。但往往快速求解也意味着做了近似和假设，达不到数学意义上的最优。在设计在线式算法时，需要考虑相关的数据服务是否能够容许近似结果与理论最优结果之间的偏差。

现有的能量管理一般假设电价是数据中心的外生变量，然而在考虑多时段的能量管理决策时，特别是数据中心可通过电储能和储冷来进行高低价差套利时，未来的电价、能价会对现有决策产生影响，此时进行价格预测就成为了必要的选择。

如何理解数据中心与其余能源网主体的互动与博弈是此类研究的难点。第一，电价受到除数据中心外的负荷（一般称为background load）的影响，这些负荷和数据中心负荷共同组成总负荷作用于电价，现有文献对于这种互相影响都处理得较为粗糙，如假设无影响或者线性变化；第二，数据中心预测电价后采取的行为一定程度上又会反过来作用于市场的价格，预测的电价并不是一个由环境决定的外生变量，而是受到数据中心自身决策影响的变化量；第三，在能源互联的背景下，由于多能源的广泛互动，多主体之间不同形式的能源价格是耦合的，如油价上涨可能导致电价上涨，对某一种的能源进行价格预测时，多能源间的耦合关系也需要进行考虑。

5.4 多级激励机制的衔接

越来越多的数据中心采用租赁式商业架构，但在租赁式架构下运营商难以对设备进行直接操作，故其在一定程度上阻碍了能量管理的应用。这种背景下，从能源互联网运营商到终端的数据中心租赁者，存在着多层次链条型的利益联结。能源互联网运营商或者服务商需要激励数据中心运营者参与广域能量市场、需求响应市场等，数据中心运营商则需要根据激励制定能量管理策略，将策略拆分为子策略后下沉到终端租赁者，激励租赁者参与其制定好的能量管理策略。

层次之间存在着博弈关系，同一层次中需要进行利益协调、分配、谈判，不同层次之间存在着激励

机制衔接的问题,跨层(直接从能源互联网运营商到终端用户)还可能存在信息不完全的问题,如何保证能量管理激励的有效性仍待研究。

5.5　数据驱动的非解析型优化

在对数据中心能耗进行建模时,大多数研究从内在物理机理或者统计拟合的角度出发,推导能耗模型的显式解析化表达式。然而很多时候这种解析化的表达仅仅在标准额定工况下有较好的近似度,在更多样的运行状况下适应程度低,使得能量管理的预期结果和实际运行情况差异较大。同时,分散化的数据中心难以用同类负荷一概论之,解析化模型参数异质程度较高,且获取方式较复杂。

为了减小解析化与实际运行情景之间的偏差,利用数据驱动的非解析型模型来建立能耗模型,并在此基础上进行数据中心能量管理的自趋优是一种新思路。其将数据中心视作黑箱智能体,在长期运行反馈数据的基础上,智能体自行学习能耗和服务器运行状态、环境温差等之间的关系,基于学习获得的能耗模式,用反馈的实时数据对能耗管理策略进行自调整。此种方法在对数据中心能耗建模的准确度方面具有优势,但其可解释性不如解析模型,未来可对具备可解释性的非解析型学习方法展开研究。

6　结论

本文从能源互联网的崭新视角出发,对数据中心能量管理的相关研究和实践进行综述,以期为数据中心能量管理的进一步研究提供参考和启发。本文既有对具体建模技术细节的阐述,也有对能量管理方法的归纳总结,且结合了实际的数据中心能量管理工程实践。相关学者未来的研究可以参考以上总结的管理方法和研究难点展开。

参考文献:

[1] 董朝阳,赵俊华,文福拴,等. 从智能电网到能源互联网:基本概念与研究框架[J]. 电力系统自动化,2014,38(15):1-11.
DONG Zhaoyang,ZHAO Junhua,WEN Fushuan,et al. From smart grid to energy internet:basic concept and research framework[J]. Automation of Electric Power Systems,2014,38(15):1-11.

[2] 曾鸣,杨雍琦,刘敦楠,等. 能源互联网"源-网-荷-储"协调优化运营模式及关键技术[J]. 电网技术,2016,40(1):114-124.
ZENG Ming,YANG Yongqi,LIU Dunnan,et al. "Generation-grid-load-storage" coordinative optimal operation mode of energy internet and key technologies[J]. Power System Technology,2016,40(1):114-124.

[3] 邓建玲. 能源互联网的概念及发展模式[J]. 电力自动化设备,2016,36(3):1-5.
DENG Jianling. Concept of energy internet and its development modes[J]. Electric Power Automation Equipment,2016,36(3):1-5.

[4] 张勇军,陈泽兴,蔡泽祥,等. 新一代信息能源系统:能源互联网[J]. 电力自动化设备,2016,36(9):1-7.
ZHANG Yongjun,CHEN Zexing,CAI Zexiang,et al. New generation of cyber-energy system:energy internet[J]. Electric Power Automation Equipment,2016,36(9):1-7.

[5] 工业和信息化部,国家机关事务管理局,国家能源局. 关于印发国家绿色数据中心试点工作方案的通知(工信部联节[2015]82 号)[EB/OL]. (2015-03-23)[2020-03-10]. http://www.miit.gov.cn/n1146295/n1652858/n1652930/n3757016/c3764169/content.html.

[6] ZHOU Z,LIU F,CHEN S,et al. A truthful and efficient incentive mechanism for demand response in green datacenters[J]. IEEE Transactions on Parallel and Distributed Systems,2020,31(1):1-15.

[7] 李丹,陈贵海,任丰原,等. 数据中心网络的研究进展与趋势[J]. 计算机学报,2014,37(2):259-274.
LI Dan,CHEN Guihai,REN Fengyuan,et al. Data center network research progress and trends[J]. Chinese Journal of Computers,2014,37(2):259-274.

[8] Cisco. Data center architecture overview:Cisco data center infrastructure 2.5 design guide[EB/OL]. (2008-05-14)[2020-03-10]. https://www.cisco.com/c/en/us/td/docs/solutions/Enterprise/Data_Center/DC_Infra2_5/DCInfra_1.html.

[9] 殷平. 数据中心研究(1):现状与问题分析[J]. 暖通空调,2016,46(8):42-53.
YIN Ping. Research on data centers(1):current situation and problem analysis[J]. Heating Ventilating & Air Conditioning,2016,46(8):42-53.

[10] TRAN N H,TRAN D H,REN S,et al. How geo-distributed data centers do demand response:a game-theoretic approach[J]. IEEE Transactions on Smart Grid,2016,7(2):937-947.

[11] 吴刚,高赐威,陈宋宋,等. 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,et al. A survey on data center power load optimization considering demand response[J]. Power System Technology,2018,42(11):3782-3788.

[12] GUO Y,LI H,PAN M. Colocation data center demand response using Nash bargaining theory[J]. IEEE Transactions on Smart Grid,2018,9(5):4017-4026.

[13] BASMADJIAN R,BOTERO J F,GIULIANI G,et al. Making data centers fit for demand response:introducing GreenSDA and GreenSLA contracts[J]. IEEE Transactions on Smart Grid,2018,9(4):3453-3464.

[14] RAHMAN A,LIU X,KONG F. A survey on geographic load balancing based data center power management in the smart grid environment[J]. IEEE Communications Surveys & Tutorials,2014,16(1):214-233.

[15] DAYARATHNA M,WEN Y,FAN R. Data center energy consumption modeling:a survey[J]. IEEE Communications Surveys & Tutorials,2016,18(1):732-794.

[16] 黄焱,王鹏,谢高辉,等. 智能电网下数据中心能耗费用优化综述[J]. 浙江大学学报(工学版),2016,50(12):2386-2399.
HUANG Yan,WANG Peng,XIE Gaohui,et al. Data center energy cost optimization in smart grid:a review[J]. Journal of Zhejiang University(Engineering Science),2016,50(12):2386-2399.

[17] 殷平. 数据中心研究(4):关键性能指标、电能使用效率PUE和EEUE[J]. 暖通空调,2017,47(4):36-45.
YIN Ping. Reasearch on data centers(4):key performance indicators,Power Usage Effectiveness(PUE) and Electric Energy Uasge Effectiveness(EEUE)[J]. Heating Ventilating & Air Conditioning,2017,47(4):36-45.

[18] JONES N. How to stop data centres from gobbling up the world's electricity[J]. Nature,2018,561(7722):163-166.

[19] CHEN M,GAO C,SONG M,et al. Internet data centers participating in demand response:a comprehensive review[J/OL]. Renewable and Sustainable Energy Reviews. (2020-01-15)[2020-

03-10]. https://www.sciencedirect.com/science/article/pii/S1364032119306744.

[20] 高赐威,曹晓峻,闫华光,等. 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 Ciwe,CAO Xiaojun,YAN Huaguang,et al. Energy management of data center and prospect for participation in demand side resource scheduling[J]. Automation of Electric Power Systems,2017,41(23):1-7.

[21] WEI T,REN S,ZHU Q. Deep reinforcement learning for joint data center and HVAC load control in distributed mixed-use buildings[J/OL]. IEEE Transactions on Sustainable Computing. (2019-04-15)[2020-03-10]. https://ieeexplore.ieee.org/document/8691496.

[22] WIERMAN A, LIU Z, LIU I, et al. Opportunities and challenges for data center demand response[C]//International Green Computing Conference. Dallas,TX,USA:IEEE,2014:1-10.

[23] LI J,BAO Z,LI Z. Modeling demand response capability by internet data centers processing batch computing jobs[J]. IEEE Transactions on Smart Grid,2015,6(2):737-747.

[24] BAHRAMI S,WONG V W S,HUANG J. Data center demand response in deregulated electricity markets[J]. IEEE Transactions on Smart Grid,2019,10(3):2820-2832.

[25] RAO L,LIU X,XIE L,et al. Minimizing electricity cost: optimization of distributed internet data centers in a multi-electricity-market environment[C]//2010 Proceedings IEEE INFOCOM. San Diego,CA,USA:IEEE,2010:1-9.

[26] CUPELLI L, SCHUTZ T, JAHANGIRI P, et al. Data center control strategy for participation in demand response programs [J]. IEEE Transactions on Industrial Informatics,2018,14(11):5087-5099.

[27] WANG W,ABDOLRASHIDI A,YU N,et al. Frequency regulation service provision in data center with computational flexibility[J/OL]. Applied Energy. (2019-10-01)[2020-03-10]. https://www.sciencedirect.com/science/article/pii/S0306261919309663.

[28] BAHRAMI S,CHEN Y C,WONG V W S. An online algorithm for data center demand response[C]//2019 53rd Annual Conference on Information Sciences and Systems(CISS). Baltimore,MD,USA:IEEE,2019:1-6.

[29] 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.

[30] GHAMKHARI M, MOHSENIAN-RAD H. Energy and performance management of green data centers: a profit maximization approach[J]. IEEE Transactions on Smart Grid,2013,4(2):1017-1025.

[31] DABBAGH M,HAMDAOUI B,RAYES A,et al. Shaving data center power demand peaks through energy storage and workload shifting control[J]. IEEE Transactions on Cloud Computing,2019,7(4):1095-1108.

[32] XU H,FENG C,LI B. Temperature aware workload management in geo-distributed data centers[J]. IEEE Transactions on Parallel and Distributed Systems,2015,26(6):1743-1753.

[33] 任华华,马克·利普尔. 燃气内燃发电机及冷热电三联供系统在数据中心的应用浅析[J]. 智能建筑电气技术,2011,5(5):57-61.
REN Huahua,MARK Lepora. Application of gas-combustion generator and CCHP system in data centers[J]. Electrical Technology of Intelligent Buildings,2011,5(5):57-61.

[34] 姜子卿,郝然,艾芊. 基于冷热电多能互补的工业园区互动机制研究[J]. 电力自动化设备,2017,37(6):260-267.
JIANG Ziqing,HAO Ran,AI Qian. Interaction mechanism of industrial park based on multi-energy complementation[J]. Electric Power Automation Equipment,2017,37(6):260-267.

[35] SHAH A,KRISHNAN N. Cycle evaluation of combined heat and power alternatives in data centers[C]//Proceedings of the 2005 IEEE International Symposium on Electronics and the Environment. New Orleans,LA,USA:IEEE,2005:19-24.

[36] ZHOU Z, LIU F, ZOU R, et al. Carbon-aware online control of geo-distributed cloud services[J]. IEEE Transactions on Parallel and Distributed Systems,2016,27(9):2506-2519.

[37] LI J, LI Z, REN K, et al. Towards optimal electric demand management for internet data centers[J]. IEEE Transactions on Smart Grid,2012,3(1):183-192.

[38] WANG P, RAO L, LIU X, et al. D-pro: dynamic data center operations with demand-responsive electricity prices in smart grid[J]. IEEE Transactions on Smart Grid,2012,3(4):1743-1754.

[39] CHEN Z,WU L,LI Z. Electric demand response management for distributed large-scale internet data centers[J]. IEEE Transactions on Smart Grid,2014,5(2):651-661.

[40] WANG H, HUANG J, LIN X, et al. Proactive demand response for data centers: a win-win solution[J]. IEEE Transactions on Smart Grid,2016,7(3):1584-1596.

[41] RAO L, LIU X, XIE L, et al. Hedging against uncertainty: a tale of internet data center operations under smart grid environment[J]. IEEE Transactions on Smart Grid, 2011, 2(3):555-563.

[42] LIN M,LIU Z,WIERMAN A,et al. Online algorithms for geographical load balancing[C]//2012 International Green Computing Conference(IGCC). San Jose, CA, USA: IEEE, 2012:1-10.

[43] YAO Y, HUANG L, SHARMA A, et al. Data centers power reduction: a two time scale approach for delay tolerant workloads[C]//2012 Proceedings IEEE INFOCOM. Orlando, FL, USA:IEEE,2012:1431-1439.

[44] CAMACHO J,ZHANG Y,CHEN M,et al. Balance your bids before your bits: the economics of geographic load-balancing [C]//Proceedings of the 5th International Conference on Future Energy Systems(e-Energy'14). New York,USA:ACM, 2014:75-85.

[45] YU L,JIANG T,CAO Y,et al. Risk-constrained operation for internet data centers in deregulated electricity markets[J]. IEEE Transactions on Parallel and Distributed Systems,2014,25(5):1306-1316.

[46] PAUL D,ZHONG W,BOSE S K. Demand response in data centers through energy-efficient scheduling and simple incentivization[J]. IEEE Systems Journal,2017,11(2):613-624.

[47] ZHAN Y,GHAMKHARI M,XU D,et al. Extending demand response to tenants in cloud data centers via non-intrusive workload flexibility pricing[J]. IEEE Transactions on Smart Grid,2018,9(4):3235-3246.

[48] XU H, LI B. Reducing electricity demand charge for data centers with partial execution[C]//Proceedings of the 5th International Conference on Future Energy Systems(e-Energy'14). New York,USA:ACM,2014:51-61.

[49] WANG K,LIN M,CIUCU F,et al. Characterizing the impact of the workload on the value of dynamic resizing in data centers[J]. ACM Sigmetrics Performance Evaluation Review,2012,40(1):405.

[50] HE Y, ELNIKETY S, LARUS J, et al. Zeta: scheduling interactive services with partial execution[C]//Proceedings of the Third ACM Symposium on Cloud Computing(SoCC'12). New York,USA:ACM,2012:1-14.

[51] LI J,QI W. Toward optimal operation of internet data center

microgrid[J]. IEEE Transactions on Smart Grid,2018,9(2):971-979.
[52] KWON S,NTAIMO L,GAUTAM N. Demand response in data centers:integration of server provisioning and power procurement[J]. IEEE Transactions on Smart Grid, 2019, 10(5):4928-4938.
[53] LU T,CHEN M,ANDREW L L H. Simple and effective dynamic provisioning for power-proportional data centers[J]. IEEE Transactions on Parallel and Distributed Systems,2013,24(6):1161-1171.
[54] BASMADJIAN R,MÜLLER L,DE MEER H. Data centres' power profile selecting policies for demand response:insights of green supply demand agreement[J]. Ad Hoc Networks,2015,25:581-594.
[55] 王乾,王彬,方建勇,等. 工业园区多能直供运营模式案例研究[J]. 电力自动化设备,2017,37(6):268-274.
WANG Qian,WANG Bin,FANG Jianyong,et al. Study of MEDS operating mode for industrial park[J]. Electric Power Automation Equipment,2017,37(6):268-274.
[56] HE T,CHONG Z R,ZHENG J,et al. LNG cold energy utilization:prospects and challenges[J]. Energy,2019,170:557-568.
[57] HUANG P,COPERTARO B,ZHANG X,et al. A review of data centers as prosumers in district energy systems:renewable energy integration and waste heat reuse for district heating[J/OL]. Applied Energy. (2020-01-15)[2020-03-10]. https://www.sciencedirect.com/science/article/pii/S0306261919317969.
[58] SAUR G,ARJONA V,CLUTTERBUCK A,et al. Hydrogen and fuel cells for data center applications project meeting:workshop report[R]. Seattle,Washington DC,USA:National Renewable Energy Lab(NREL),2019.
[59] IT之家. 贵州今年建成华为数据中心、iCloud数据中心等项目[EB/OL]. (2020-01-17)[2020-03-10]. https://tech.163.com/20/0117/08/F332775N000999LD.html.
[60] 阿里技术保障. 阿里巴巴千岛湖数据中心正式启动[EB/OL]. (2015-09-08)[2020-03-10]. http://dc.idcquan.com/jfjs/77276.shtml.
[61] 国家发展改革委,科技部,工业和信息化部,等. 关于印发《加快推进天然气利用的意见》的通知[EB/OL]. (2017-07-04)[2020-03-10]. http://www.gov.cn/xinwen/2017-07/04/content_5207958.htm.
[62] 殷平. 数据中心研究(5):燃气冷热电三联供[J]. 暖通空调,2017,47(6):1-8.
YIN Pin. Research on data centers(5):gas CCHP[J]. Heating Ventilating & Air Conditioning,2017,47(6):1-8.
[63] 腾讯. 探寻腾讯云基石,走进腾讯上海青浦数据中心[EB/OL]. (2018-03-16)[2020-03-10]. https://cloud.tencent.com/developer/article/1063962.
[64] 中国南方电网有限责任公司. 全国最大地下天然气分布式能源站建成[EB/OL]. (2017-03-22)[2020-03-10]. https://www.nengapp.com/news/detail/892442.
[65] 新浪科技. Facebook投资1亿美元在丹麦建立数据中心[EB/OL]. (2017-01-20)[2020-03-10]. http://tech.sina.com.cn/i/2017-01-20/doc-ifxzuswq2799674.shtml.
[66] EBay. 6 MW bloom fuel cell installation to provide prime power for eBay's new data centre[EB/OL]. (2012-06-21)[2020-03-10]. http://www.fuelcelltoday.com/news-archive/2012/june/6-mw-bloom-fuel-cell-installation-to-provide-prime-power-for-ebay's-new-data-centre.

作者简介:

冯　成

冯　成(1997—),男,四川成都人,博士研究生,主要研究方向为信息物理融合、电力大数据(**E-mail**:fengc19@mails.tsinghua.edu.cn);

王　毅(1992—),男,湖北仙桃人,博士后,主要研究方向为电力大数据、综合能源系统、信息物理融合(**E-mail**:yiwang@eeh.ee.ethz.ch);

陈启鑫(1982—),男,广东潮州人,副教授,博士,通信作者,主要研究方向为低碳电力技术、电力大数据、能源互联网、电力市场、环境调度等(**E-mail**:qxchen@tsinghua.edu.cn)。

(编辑　李玮)

Review of energy management for data centers in energy internet

FENG Cheng[1],WANG Yi[2],CHEN Qixin[1],ZHANG Xuan[3],LUO Gang[3]
(1. State Key Laboratory of Control and Simulation of Power Systems and Generation Equipment,Tsinghua University,Beijing 100084,China;2. ETH Zurich,Zurich 8092,Switzerland;
3. Power Dispatch and Control Centre,Guangdong Power Grid Co.,Ltd.,Guangzhou 510699,China)

Abstract:Data centers play an important role in the big data era. With the continuous expansion of data centers,the problem of their huge energy consumption has attracted a lot of attention and thus the energy management has become a major focus for data center industry. In the energy internet environment,data centers are expected to have more diversified choices to gain energy supply,discover more flexible and efficient energy management methods,and their energy management modes that only rely on electricity could be reformed fundamentally. From a new perspective of energy internet,the research and practice related to the energy management for data centers are reviewed. Firstly,the service structure and the infrastructure of data centers are discussed. Then,the energy consumption structure,energy efficiency and energy consumption model of data centers are provided. Afterwards,the energy management methods of data centers in the context of energy internet are analyzed in detail. At last,the challenges and difficulties of energy management for data centers in the context of energy internet in the future are discussed.

Key words:data center;energy internet;energy management;energy consumption model;demand response