

计及数据中心负载特性的多时间尺度优化调度

帅文利¹, 杨 秀¹, 赵晓莉¹, 孙改平¹, 李莉华²

(1. 上海电力大学 电气工程学院, 上海 200090; 2. 国网上海市电力公司电力科学研究院, 上海 200092)

摘 要: 为降低数据中心能耗, 提出一种基于负载特性的新型多时间尺度调度模型。在日前阶段, 通过分析批处理负载的到达率、可延迟时间、执行时间及内部 DAG 结构等底层逻辑特性, 建立矩阵模型参与调度优化, 以平滑服务器能耗曲线; 在日内阶段, 结合冗余备份与异地转移策略应对负载波动, 构建滚动优化模型, 实时调整日内计划。采用 YALMIP + CPLEX 求解器进行仿真验证, 结果表明该模型显著降低了数据中心的运行成本, 对提升能效管理水平具有重要实践价值。

关键词: 数据中心; 负载特性; 底层逻辑; 矩阵模型; 多时间尺度

中图分类号: TM732

文献标志码: A

DOI: 10.20204/j.sp.2025.07001

Multi-timescale Optimal Scheduling Considering Data Center Load Characteristics

SHUAI Wenli¹, YANG Xiu¹, ZHAO Xiaoli¹, SUN Gaiping¹, LI Lihua²

(1. School of Electrical Power Engineering, Shanghai University of Electric Power, Shanghai 200090, China;

2. State Grid Shanghai Electric Power Research Institute, Shanghai 200092, China)

Abstract: To reduce data center energy consumption, this study proposes a novel multi-timescale scheduling model based on load characteristics. In the day-ahead phase, by analyzing the underlying logical characteristics of batch workloads—including arrival rate, delay tolerance duration, execution time, and internal DAG (Directed Acyclic Graph) structure—a matrix model is established to participate in scheduling optimization, thereby smoothing server energy consumption curves. In the intraday phase, a rolling optimization model is constructed to dynamically adjust intraday plans by integrating redundancy backup and geo-migration strategies to address load fluctuations. The model is validated using the YALMIP + CPLEX solver. Results demonstrate that the proposed model significantly reduces the operational costs of data centers and offers substantial practical value for enhancing energy efficiency management capabilities.

Key words: data center; load characteristics; underlying logic; matrix model; multiple-timescale

0 引言

社会对计算资源的需求激增。据测算, 我国数据量全球占比将从 2018 年的 23.4% 升至 2025 年的 27.8%, 数据中心能耗体量激增^[1]。因此急需挖掘数据中心负载调节潜力, 降低数据中心日益增长的能耗成本。数据中心工作负载可根据对响应时间要求的不同划分为交互式负载和批处理式负载^[2]。交互式负载数据量较小, 具有空间灵活性^[3-4], 但对时延敏感且受用户行为影响, 计算需求波动较大; 批处理负载数据量较大, 对时效性要求不高, 具

有时间灵活性以及较强的稳定性和周期性^[5-6]。为提高数据中心资源利用率, 通常对交互式负载和批处理负载进行混合部署^[7-8]。

对于批处理负载, 其内部复杂的组成对负载调度的灵活性造成影响, 所以需要考虑负载底层逻辑。现有研究主要关注到达率、延迟时间、执行时间等负载特性的差异^[9-10]。文献[11]在延迟时间上限内, 利用不同批处理负载时间灵活性优化算力需求处理策略。文献[12]批处理负载根据处理周期分类, 采用队列模型进行优化调度。文献[13]将批处理负载划分为规模大小不同的两类, 并考虑其可调节能力, 使得系统经济性和灵活性大幅提高。但是这些研究都忽略了负载内部结构对调度的影响。批处理负载由各种批处理作业(job)组成, 通常一个批处理作业又被分为多个任务(task), 任务间数据依赖关系用有

基金项目: 国家自然科学基金资助项目(52177098)

Project Supported by the National Natural Science Foundation of China (52177098)

向无环图(Directed Acyclic Graph, DAG)表示^[14]。文献[15]进一步分析负载的特征,以阿里巴巴数据中心为例,对批处理负载的 DAG 结构进行了分类和总结。文献[16]描绘了批处理作业的 DAG 结构以及任务关键路径,计算了任务等待时间,作业间隙时间等指标,指出其可能的任务调度优化空间。但是这些文献都是对批处理负载的部分特征进行建模分析,针对批处理负载特性的建模以及与调度的结合工作仍存在不足^[17]。

对于交互式负载,其不确定性与突发性^[18],以及可再生能源的不确定性,增加了数据中心调度运行的复杂性^[19-20]。合理调度工作负载能促进可再生能源的消纳^[21-23],但是单一时间尺度的调度计划难以实现数据中心资源调度的经济性最优^[24]。文献[25]建立数据中心两阶段调度模型,利用冗余计算资源,有效解决工作负载的不确定性问题,文献[26]日内对交互式负载进行空间调度,通过日前-日内两阶段协同调度减少了预测误差所造成的经济损失,但是忽略了批处理负载内部结构带来的调度潜力以及混部环境下工作负载间资源竞争和性能干扰^[27-28]。

基于此,本文在日前阶段保证交互式负载运行的情况下,设置“冗余备份”后对批处理负载进行调度,考虑负载底层逻辑特性,构建批处理负载矩阵模型,并加入延迟量约束、DAG 约束以及用户满意度。日内阶段,建立滚动优化模型,交互式负载通过异地调度以及服务器备份容量保证其即时响应,同时各设备协调出力减少由日前预测误差造成的经济损失。

1 数据中心负载及能耗模型

1.1 数据中心结构

数据中心作为信息技术的关键基础设施,其架构包含多个相互关联的部分,数据中心供电系统整合风电、光伏等可再生能源,并配备燃气轮机和储能设备,通过与大电网互联实现电力平衡。服务器集群是数据中心处理数据的核心组件,它响应来自数据用户集群的批处理负载和交互式负载请求。到达数据中心的工作负载组成复杂,不同类型的工作负载对服务器资源的需求和使用模式存在显著差异,从而直接影响服务器能耗,而数据中心的主要耗电设备为服务器和冷却服务器散热的空调,因此建立精细化的工作负载模型并制定合理的负载调度策略对于降低数据中心能耗至关重要。本文

研究的数据中心结构如图 1 所示。

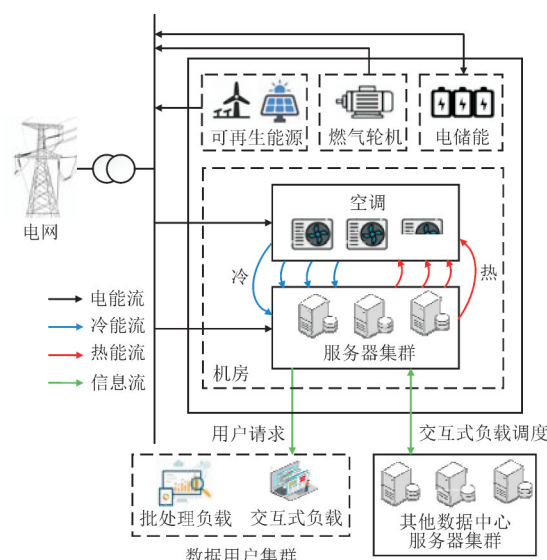


图1 数据中心结构示意图

Fig.1 Schematic diagram of data center structure

1.2 批处理负载模型

在批处理系统中,‘job’代表了一个完整的作业或任务集合,而‘task’则是作业中的一个具体执行单元,task 之间具有数据依赖关系。单任务作业的可延迟时间和执行时间等同于该任务的时间,多任务作业的执行时间为内部 DAG 最长路径时间,并且其可延迟时间是对于整体来说的。考虑负载内部 DAG 结构可以优化任务执行顺序,充分利用其存在的任务调度优化空间。批处理负载通常是按照预定计划执行,负载到达以及执行存在周期性,因此在本文建立批处理负载模型时,考虑在调度周期内能处理上一周期的延迟批处理负载,同时也能将当日未处理的负载转移至下一周期处理。

1.2.1 负载调度模型

假设到达数据中心的批处理负载根据其负载执行时间,负载可延迟处理时间以及负载内部 DAG 结构等特点可以聚为 M 类 job,包括 $job_1, job_2, \dots, job_M$,其中 job_m 表示第 m 类 job 且 $m \in \{1, \dots, M\}$,将 M 类 job 展开到 task 层,总共有 K 类 task,包括 $task_1, task_2, \dots, task_K$,其中 $task_k$ 表示第 k 类 task 且 $k \in \{1, \dots, K\}$,构建 task 集合,总共有 M 个 task 集合,同一集合的 task 具有与 job 相同的到达率且在 job 层的可调度域内进行调度, T 为调度周期。 job_m 以及其包含 task 的可调度时间、执行时间和可延迟处理时间之间的关系为:

$$E_{DW}^{job,m} = E_{solve}^{job,m} + E_{delay}^{job,m} \quad (1)$$

$$E_{solve}^{task,k} + E_{delay}^{job,m} \leq E_{DW}^{job,m} \quad (2)$$

$$E_{delay}^{task,k} = E_{DW}^{job,m} - E_{solve}^{task,k} \quad (3)$$

式中: $E_{DW}^{job,m}$, $E_{solve}^{job,m}$, $E_{delay}^{job,m}$ 分别为 job_m 的可调度时间、执行时间、可延迟处理时间; $E_{solve}^{task,k}$ 为 $task_k$ 的执行时间且此 $task$ 属于第 m 个 job ; $E_{delay}^{task,k}$ 为 $task_k$ 可延迟处理时间段。

一个优化周期内各时段数据中心的 $task$ 到达率矩阵 $Y_{K \times T}^{task}$ 如式(4)所示:

$$Y_{K \times T}^{task} = (y_{k,t}^{task}), k \in \{1, \dots, K\}, t \in \{1, \dots, T\} \quad (4)$$

式中: $y_{k,t}^{task}$ 为 t 时刻到达的第 k 类 $task$ 到达率。

对到达数据中心的 $task$ 都建立大小为 $T \times T$ 的矩阵, $task_k$ 的负载转移矩阵 $X_{T \times T}^{task,k}$ 、开始执行矩阵 $G_{T \times T}^{task,k}$ 和完成矩阵 $R_{T \times T}^{task,k}$ 分别如式(5)、式(6)和式(7)所示:

$$X_{T \times T}^{task,k} = (x_{t,j}^{task,k}), t \in \{1, \dots, T\}, j \in \{1, \dots, T\} \quad (5)$$

$$G_{T \times T}^{task,k} = X_{T \times T}^{task,k} \cdot F_{T \times T}^k \quad (6)$$

$$R_{T \times T}^{task,k} = Mod(G_{T \times T}^{task,k} - (E_{solve}^{task,k} - 1)) \quad (7)$$

开始执行矩阵约束为:

$$G_{T \times T}^{task,k} \geq 0 \quad (8)$$

$$\sum_{j=1}^T G_{t,j}^{task,k} = y_{k,t}^{task} \quad (9)$$

式中: $x_{t,j}^{task,k}$ 为 t 时刻到达的第 k 类 $task$ 转移到 j 时刻开始进行处理的负载量; $F_{T \times T}^k$ 为第 k 类负载的 0-1 执行约束矩阵, 在 $E_{delay}^{task,k}$ 时间段内为 1, 负载可执行, 其余时段为 0, 不可执行; Mod 为矩阵平移函数。

服务器对 $task_k$ 的处理量矩阵 $D_{T \times T}^{task,k}$ 如式(10)所示:

$$D_{T \times T}^{task,k} = \sum_{i=0}^{E_{delay}^{task,k}-1} Mod(G_{T \times T}^{task,k} - i) \quad (10)$$

j 时刻数据中心的批处理负载处理量 D_j^{DW} 如式(11)所示:

$$D_j^{DW} = \sum_{t=1}^T \sum_{k=1}^K D_{t,j}^{task,k} \quad (11)$$

式中: $D_{t,j}^{task,k}$ 为服务器对 t 时刻到达的 $task_k$ 在 j 时刻的处理量。

1.2.2 延迟量约束

与 CPU 利用率相比, 大多数服务器具有更高的内存利用率, 内存正在成为影响工作负载性能的瓶颈^[29]。同一时刻数据中心的负载延迟量越多, 内存占用越高, 从而影响系统内资源利用率, 所以需要合理限制负载延迟量, t 时刻负载的延迟量 V_t^{DW} 如式(12)所示:

$$V_t^{DW} = V_{t-1}^{DW} + y_{k,t}^{task} \cdot E_{solve}^{task,k} - D_t^{DW} \quad (12)$$

延迟量约束为:

$$V_t^{DW} \leq \psi_1 \quad (13)$$

式中: V_{t-1}^{DW} 为 $t-1$ 时刻批处理负载延迟量; D_t^{DW} 为 t

时刻批处理负载处理量; ψ_1 为最大允许负载延迟量。

1.2.3 DAG 约束

到达数据中心的不同 job 内部有不同的 DAG 结构, 而 DAG 关系的实质就是对 job 内部的 $task$ 之间处理顺序进行优先级排序。根据所有 job 包含的 $task$ 间的独立以及依赖关系构建 0-1 依赖矩阵 $A_{K \times K}$, $p, q \in \{1, \dots, K\}$, $A_{p,q} = 0$ 表示 $task_p$ 与 $task_q$ 具有独立执行关系, $A_{p,q} = 1$ 代表 $task_p$ 与 $task_q$ 有直接依赖关系, $task_p$ 为先驱任务, $task_q$ 在 $task_p$ 执行完毕之后再开始执行, 此时对应的 $A_{q,p} = -1$ 。当存在依赖关系, 即 $A_{p,q} = 1$ 时, t 时刻到达的任务在 j 时刻会处于 3 种可能的处理情况之一。

1) t 时刻到达的先驱任务 $task_p$ 在 j 时刻有部分处理完成, 如式(14)所示:

$$\begin{cases} R_{t,j}^{task,p} > 0, R_{t,j}^{task,q} = 0 \\ \sum_{i=t}^j G_{t,i}^{task,q} = 0 \end{cases} \quad (14)$$

2) t 时刻到达的后执行任务 $task_q$ 在 j 时刻有部分处理完成, 如式(15)所示:

$$R_{t,j}^{task,q} > 0, R_{t,j}^{task,p} = 0 \quad (15)$$

3) t 时刻到达的 $task_p$ 和 $task_q$ 在 j 时刻都无任务完成, 如式(16)所示:

$$R_{t,j}^{task,p} = 0, R_{t,j}^{task,q} = 0 \quad (16)$$

当 job 中包含的所有 $task$ 处理完成时才判定此 job 处理完成, 所以还需对其中的 $task$ 进行分析, 同一类 job 中的 $task$ 存在 3 种可能的状态^[16]:

1) 无 DAG 关系。 $task_p$ 不作为任意 $task$ 的先驱任务, 可以独立执行, 不受其他 $task$ 执行的影响, 如式(17)所示:

$$\forall p \in [1, K], A_{p,q} = 0 \quad (17)$$

2) 有 DAG 关系无依赖。 $task_p$ 本身不依赖于任何其它 $task$, 并且为其所属 job 包含的其余 $task$ 的依赖。

3) 有 DAG 关系有依赖。 $task_p$ 有 1 个或者多个前驱依赖, 如式(18)所示:

$$\exists p \in [1, K], A_{p,q} = -1 \quad (18)$$

当 $task$ 存在前驱依赖但其本身并不作为其他 $task$ 的前驱任务时, 认为此 $task$ 为末端任务, 如式(19)所示:

$$\begin{cases} \forall p \in [1, K], A_{p,q} \leq 0 \\ \exists p \in [1, K], A_{p,q} = -1 \end{cases} \quad (19)$$

通过对比 job 中的末端 $task$ 与独立执行 $task$ 的完成情况, 可直接评估其完成情况。

1.2.4 满意度模型

对于批处理作业, 用户通常更关心整个作业完

成所需的时间,而不是单个任务的执行情况。在将对批处理作业的调度从 job 层细化到 task 层后,实际优化调度在 task 层发生,而用户满意度的感知变化则是在 job 层上发生。不同的 job,由于大小以及优先度的不同,其与用户签订的协议中可延迟时间不同,延迟相同的时间对用户来说满意度变化不同,建立对应 job 类型的满意度矩阵 $\mathbf{W}_{T \times T}^m$ 如式(20)所示:

$$\mathbf{W}_{t,t+i}^m = \begin{cases} z1^m - (z1^m - z2^m) \times \frac{(i - E_{\text{solve}}^{\text{job},m})}{E_{\text{delay}}^{\text{job},m}}, E_{\text{solve}}^{\text{job},m} < i \leq E_{\text{DW}}^{\text{job},m} \\ 0, \text{其他} \end{cases} \quad (20)$$

式中: $\mathbf{W}_{t,t+i}^m$ 为 t 时刻到达的负载在 $t+i$ 时刻完成时的满意度系数; $z1^m$ 为 job_m 到达数据中心后立即被执行所对应的满意度系数; $z2^m$ 为 job_m 在用户规定的截止时间完成时所对应的用户满意度系数。

假设在 job_m 中 task_{k1} , task_{k2} 都为末端任务, task_{k3} 为独立执行可并行任务, job_m 的完成矩阵 $\mathbf{H}_{T \times T}^m$ 建立过程如式(21)、式(22)所示:

$$S_{t,j}^m = \min \left(\sum_{i=t}^j R_{t,i}^{\text{task},k1}, \sum_{i=t}^j R_{t,i}^{\text{task},k2}, \sum_{i=t}^j R_{t,i}^{\text{task},k3} \right) \quad (21)$$

$$\begin{cases} H_{t,j}^m = S_{t,j}^m, j = t \\ H_{t,j}^m = S_{t,j}^m - S_{t,j-1}^m, j > t \end{cases} \quad (22)$$

式中: $S_{t,j}^m$ 为 t 时刻到达的三类任务在 j 时刻的最小完成量。

由此,对 job_m 进行调度后的用户满意度矩阵 $\mathbf{B}_{T \times T}^m$ 如式(23)所示:

$$\mathbf{B}_{T \times T}^m = \mathbf{W}_{T \times T}^m \cdot \mathbf{H}_{T \times T}^m \quad (23)$$

1.3 交互式负载模型

交互式负载对延迟敏感,在到达数据中心后需要立即执行,因此优先考虑集中资源处理交互式负载以保证其即时响应要求。 t 时刻数据中心的交互式负载处理量 D_t^{AW} 如式(24)所示:

$$D_t^{\text{AW}} = y_t^{\text{AW}} \quad (24)$$

式中: y_t^{AW} 为交互负载在 t 时刻的到达率。

混部环境下,批处理负载与交互式负载资源抢占会使负载的运行环境更加复杂,导致交互式负载资源配置失败并严重损害批处理作业的性能。为了避免工作负载之间的性能干扰,让部分服务器处于待机或者关闭的状态,批处理负载无法使用这部分资源,在既定资源下进行调度,由此设置“冗余备份” D_t^{DW} 如式(25)所示:

$$D_t^{\text{DW}} \leq (1 - \vartheta^{\text{ser}}) \cdot n^{\text{ser}} \cdot \mu^{\text{ser}} - D_t^{\text{AW}} \quad (25)$$

式中: μ^{ser} 为单台服务器的服务率,即单位时间内可处理负荷量; n^{ser} 为数据中心服务器配置总数; ϑ^{ser}

为服务器冗余备用系数。

交互式负载可以快速异地调度,当 t 时刻将负载调度到其他数据中心进行处理时,数据中心的交互式负载处理量 D_t^{AW} 如式(26)所示,调度约束 Δy_t^{AW} 如式(27)所示:

$$D_t^{\text{AW}} = y_t^{\text{AW}} - \Delta y_t^{\text{AW}} \quad (26)$$

$$\Delta y_t^{\text{AW}} \leq \Delta y_t^{\text{MAX}} \quad (27)$$

式中: Δy_t^{AW} 为 t 时刻异地调度负载量; Δy_t^{MAX} 为 t 时刻允许的最大负载异地调度量。

1.4 数据中心能耗模型

1.4.1 服务器能耗模型

服务器的能耗主要取决于 CPU 利用率、服务器功率参数及当前时刻开机数量。本文假设所有服务器的型号相同,到达数据中心的负载将平均分配到各开启服务器,故开启的服务器 CPU 利用率相同。 t 时段所有服务器的总能耗如式(28)所示:

$$P_t^{\text{IT}} = [(P^{\text{peak}} - P^{\text{idle}}) \cdot u_t + P^{\text{idle}}] \cdot n_t^{\text{on}} + P^{\text{sleep}} \cdot n_t^{\text{sleep}} \quad (28)$$

式中: P_t^{IT} 为 t 时段服务器总运行功率; u_t 为 t 时段单台服务器的 CPU 利用率; P^{peak} , P^{idle} 分别为单台服务器的峰值功率和静默功率; P^{sleep} 为单台服务器的待机功率; n_t^{sleep} , n_t^{on} 分别为 t 时段服务器的待机数量和开机数量。

CPU 利用率由服务器正在处理的工作负载量、服务器开机数量和服务率决定,包括正在处理的批处理负载以及交互式负载。根据式(11)、式(24)可得:

$$D_t = D_t^{\text{AW}} + D_t^{\text{DW}} \quad (29)$$

$$u_t = D_t / (n_t^{\text{on}} \mu^{\text{ser}}) \quad (30)$$

服务器约束为:

$$0 \leq u_t \leq U^{\text{max}} \quad (31)$$

$$n^{\text{ser}} = n_t^{\text{on}} + n_t^{\text{sleep}} + n_t^{\text{close}} \quad (32)$$

式中: D_t 为 t 时段所有服务器的负载处理总量; U^{max} 为 CPU 利用率上限; n_t^{close} 为 t 时段服务器的关机数量。

1.4.2 空调能耗模型

服务器运行时会产生大量热能使得温度升高,过高的温度会影响服务器的稳定运行甚至导致设备损坏,因此数据中心配置冷却设备来调节机房内温度,冷却设备运行产生的冷量中用于吸收 IT 设备和环境热量的部分约占空调设备冷负荷的 90%^[30],空调的冷负荷来源如式(33)一式(35)所示:

$$Q_t^{\text{air}} = Q_t^{\text{IT}} + Q_{\text{env}} \quad (33)$$

$$Q_t^{\text{IT}} = 0.97 P_t^{\text{IT}} \quad (34)$$

$$Q_{\text{env}} = -e_1 A (T_t^{\text{in}} - T_t^{\text{out}}) \quad (35)$$

式中: Q_t^{air} 为 t 时段空调的冷负荷; Q_t^{IT} 为 t 时段 IT 设备

产生的热能(0.97为散热比); Q_{env} 为环境热量,与机房内外温度差有关; e_1 为机房围护结构传热系数; A 为机房传热面积; T_i^{out} 为机房外部温度; T_i^{in} 为机房内部温度。

空调设备的能耗如式(36)所示:

$$P_i^{\text{air}} = \frac{Q_i^{\text{air}}}{\theta_i^{\text{cop}}} = \frac{0.97P_i^{\text{IT}} - e_1 A (T_i^{\text{in}} - T_i^{\text{out}})}{\theta_i^{\text{cop}}} \quad (36)$$

式中: θ_i^{cop} 为冷负荷和空调能耗的比值性能系数,即空调的制冷效率。

根据式(28)、式(36)可得数据中心总能耗 P_i^{tol} 为:

$$P_i^{\text{tol}} = P_i^{\text{IT}} + P_i^{\text{air}} \quad (37)$$

2 两阶段数据中心优化调度模型

2.1 调度框架及求解流程

本文的日前日内两阶段优化调度模型如图2所示。

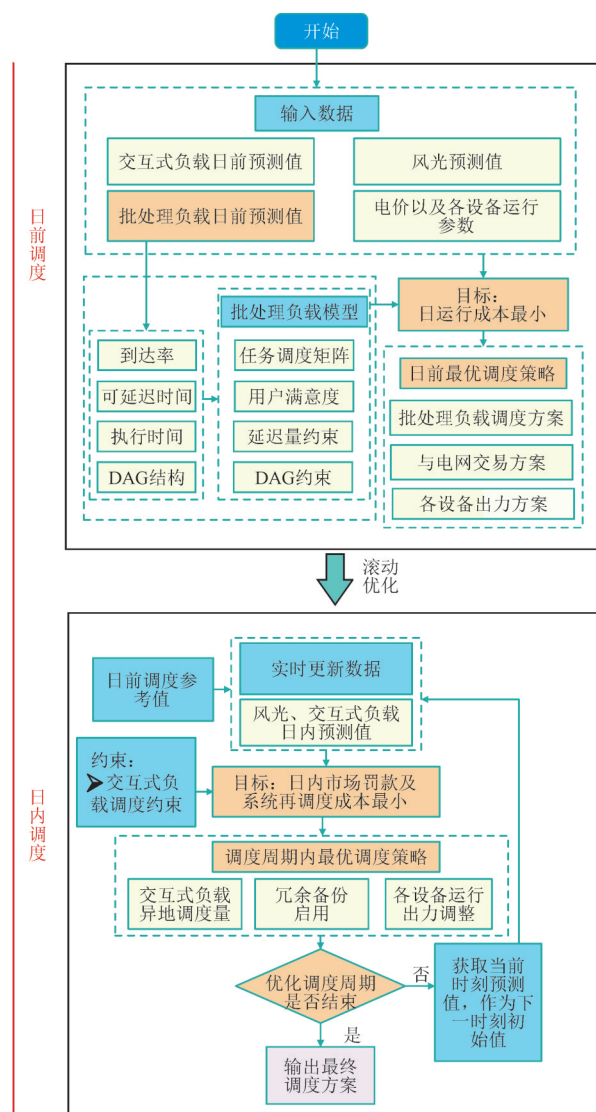


图2 数据中心两阶段优化决策模型求解流程

Fig.2 Solution process for two-stage optimization decision model in data centers

2.2 日前长时间尺度优化模型

在日前阶段以数据中心运行成本最低为目标,考虑功率平衡、燃气轮机、储能以及工作负载调度等约束条件,优化求解次日时段 T 各数据中心的最优调度计划。其目标函数如式(38)所示:

$$\min C_{\text{DA}} = \sum_{t=1}^T (C_t^{\text{EX}} + C_t^{\text{GT}} + C_t^{\text{ES}} + C_t^{\text{DW}} + C_t^{\text{RES}} + C_t^{\text{IT}}) \quad (38)$$

式中: C_{DA} 为数据中心总运行成本; C_t^{EX} 为电网交互成本; C_t^{GT} 为可控机组运行成本; C_t^{ES} 为储能运行维护成本; C_t^{DW} 为批处理负载调度补偿成本; C_t^{RES} 为可再生能源运维成本; C_t^{IT} 为动态服务器开销成本。每部分的具体表达式为:

$$C_t^{\text{EX}} = \gamma_t^1 P_t^{\text{net}} = \gamma_t^1 (P_t^{\text{buy}} - P_t^{\text{sell}}) \quad (39)$$

式中: γ_t^1 为日前电网电价; $P_t^{\text{buy}}, P_t^{\text{sell}}$ 分别为在 t 时刻电网交互的购电和售电功率; P_t^{net} 为 t 时刻与电网交互功率。

$$C_t^{\text{RES}} = \sum_{i=1}^{24} c_{\text{pv}} P_t^{\text{pv}} + c_{\text{pw}} P_t^{\text{pw}} \quad (40)$$

式中: $c_{\text{pv}}, c_{\text{pw}}$ 为光伏和风机运维成本系数, $P_t^{\text{pv}}, P_t^{\text{pw}}$ 为 t 时刻光伏和风机提供的功率。

$$C_t^{\text{GT}} = a(P_t^{\text{GT}})^2 + bP_t^{\text{GT}} + c + \delta_t^u C_u^{\text{GT}} + \delta_t^d C_d^{\text{GT}} \quad (41)$$

式中:本文中可控机组主要为燃气机组, a, b, c 分别为燃气机组的二次项、一次项和常数项发电成本系数; P_t^{GT} 为数据中心 t 时段燃气机组的输出功率; δ_t^u, δ_t^d 为布尔变量,分别表示燃气机组在 t 时段是否启动、停止,启动则置1,否则置0; $C_u^{\text{GT}}, C_d^{\text{GT}}$ 分别为燃气机组的启动和停止成本。

$$C_t^{\text{ES}} = \lambda^{\text{ES}} (P_t^{\text{cha}} + P_t^{\text{dis}}) \quad (42)$$

式中: λ^{ES} 为单位储能运行维护成本; $P_t^{\text{cha}}, P_t^{\text{dis}}$ 分别为储能在 t 时段的充电功率和放电功率。

$$C_t^{\text{IT}} = \beta_1 |n_t^{\text{close}} - n_{t-1}^{\text{close}}| \quad (43)$$

式中: β_1 为单位服务器开关成本。

从与用户的长期合作发展来看,不能只关注数据中心本身的经济利益,若为了达到最高的经济效益而忽略了用户的满意度本身来说不利于数据中心的长期发展,所以在此考虑对调度造成的用户满意度变化进行补偿。

$$C_t^{\text{DW}} = \xi^{\text{DW}} \cdot \sum_{j=1}^T \sum_{m=1}^M B_{j,t}^m \quad (44)$$

式中: ξ^{DW} 为满意度补偿系数; $B_{j,t}^m$ 为在 j 时刻到达 job_m 在 t 时刻完成的满意度。

数据中心运行时需要满足电功率平衡,即式(38)的约束为:

$$P_t^{pv} + P_t^{pw} + P_t^{dis} + P_t^{net} + P_t^{GT} = P_t^{cha} + P_t^{tol} \quad (45)$$

2.3 日内短时间尺度优化模型

日内阶段的调度策略是在日前阶段的基础上,将日前批处理负载处理计划作为确定值代入日内调度中,考虑预测误差、环境等因素导致的调度计划与实际运行之间的偏差,依据更新后的风电、光伏出力以及交互式预测数据进行日内调整,尽可能跟踪日前计划以减少日内市场的罚款,日内阶段的目标函数为:

$$\min \sum_{t=t_0}^{\Delta T} (C_t^{PU} + C_t^{AW} + C_t^{GT} + C_t^{ES} + C_t^{RES} + C_t^{IT} + C_t^{DW}) \quad (46)$$

式中: t_0 为当前调度时段的初始时刻; ΔT 为日内调度1个调度周期时长; C_t^{PU} 为日内功率偏差惩罚成本; C_t^{AW} 为日内交互式负载调度补偿成本。

日内对超出日前购电计划用电量以高于实时电价的价格进行罚款,对低于日前购电计划的量以低于实时电价的价格进行罚款,如式(47)所示:

$$C_t^{PU} = \begin{cases} \varepsilon_1 \gamma_t^2 \Delta P_t^{net}, & \Delta P_t^{net} > 0 \\ \varepsilon_2 \gamma_t^2 \Delta P_t^{net}, & \Delta P_t^{net} < 0 \end{cases} \quad (47)$$

式中: $\varepsilon_1, \varepsilon_2$ 分别为高于、低于日前电能交易计划时的惩罚系数; γ_t^2 为日内电网电价; ΔP_t^{net} 为日内购电计划与日前计划的偏差。

$$C_t^{AW} = \xi^{AW} \cdot \Delta y_t^{AW} \quad (48)$$

式中: ξ^{AW} 为单位交互负载调度补偿费用。

3 算例分析

3.1 参数设置

数据中心设置8000台同构服务器,服务器的参数满载能耗为600W,空载能耗为200W,休眠能耗为50W,平均服务率为20个/s。日前为24h出力方案,1h为时间粒度,日内6h预测时域,15min调度时域,日前日内预测的交互式负到达曲线、批处理负载的到达曲线、日前日内预测的交互式负到达曲线以及风光出力曲线数据如图3中所示。

由于数据中心运行需要保证供电的稳定性和可靠性,其内部不稳定的可再生能源用量不宜过大,因此本文设置的光伏风电出力较小,以保证可再生能源在不确定性情况下,预测出力与实际出力相差不会过大从而影响数据中心用电。

对于到达的批处理负载,本文设置4个job类型,其内部DAG结构如图4所示。

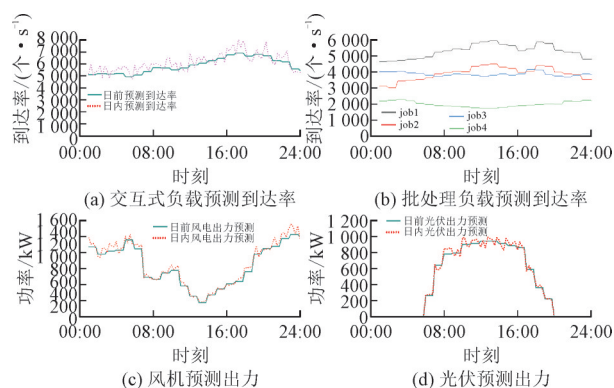


图3 算例数据

Fig.3 Example dataset

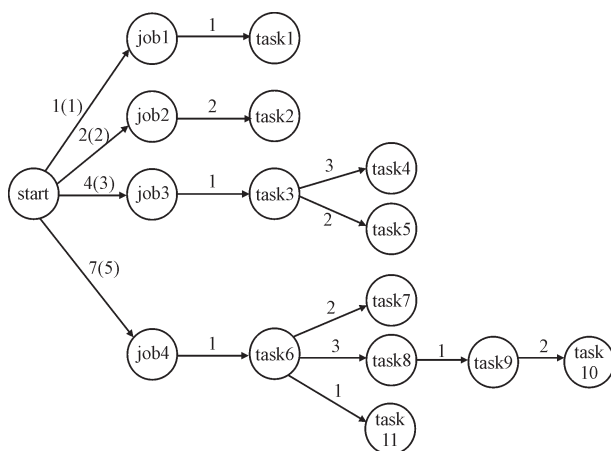


图4 批处理负载特征及DAG结构

Fig.4 Batch workload features and DAG structure

图4指向job的箭头上标签中,括号外的数值表示job的最短执行路径时间,括号内的数值表示延迟容忍时间;指向task的数值表示task的执行时间。

本文在MATLAB2021a平台上进行仿真计算,通过调用Yalmip/CPELX工具箱求解所涉及的优化问题。

3.2 日前优化调度分析

3.2.1 调度结果分析

日前数据中心可再生能源出力、各设备出力曲线,空调、服务器能耗优化调度后结果如图5所示。

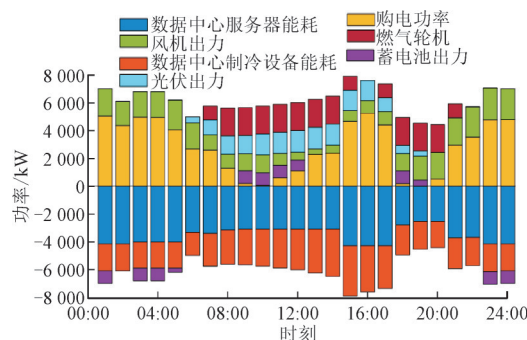


图5 日前调度功率平衡图

Fig.5 Day-ahead scheduling power balance

根据图 5 可以看出,日前优化调度后的购电量曲线呈现“双峰双谷”的趋势,这是由于 09:00—14:00 和 18:00—20:00 属于用电高峰,此时电价较高。经过优化之后,在午间和晚间电价高峰时,光伏风电以及燃气轮机出力,填补了部分数据中心电力消耗,在电价较低时段储存起来的部分电量也在此时释放,批处理负载调度到电价较低时段进行处理。午间电价高峰的部分负载延迟到 14:00—18:00 进行处理,同时为降低晚间电价高峰时段服务器耗电,此时段开启大量服务器处理负载以减少延迟到晚高峰处理的负载。22:00—6:00 处于电价低谷,并且此时服务器资源充足,因此会有较多批处理负载延迟到这段进行处理。

日前优化调度后,不同类批处理负载处理完成情况如图 6 所示。

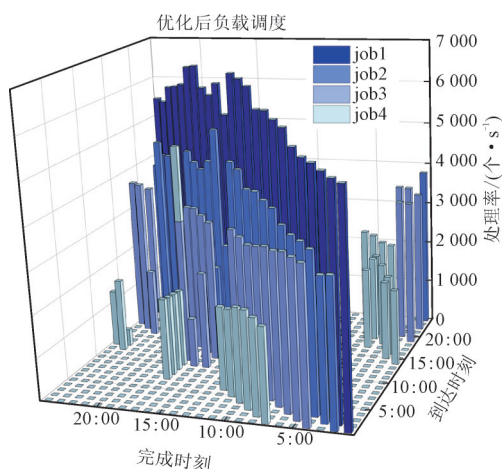


图6 优化后批处理负载完成情况

Fig.6 Completion status of optimized batch workloads

从图 6 中可以看出,调度优化量 job4 最大,其负载完成时段最为集中,以避免电价高峰期。job1 的调度优化量最小,大部分负载都在到达后立即进行处理。所有负载调度灵活性 $job4 > job3 > job2 > job1$,即负载量更大的批处理负载时间灵活性更好,调度时优先调用负载量更大的批处理负载,这是由于其容忍时间相对较长,可调度范围更广,延迟调度时对用户满意度影响更小,需要的补偿费用更少。同时保证了其内部任务执行顺序之下,充分利用时间间隙,使得其有了任务调度优化空间,将小粒度的任务调度到电价较低时段进行处理。负载量较小的批处理负载,其能为数据中心的能耗优化起一定作用,但由于与用户签订的协议中,可延迟时间范围小,用户对其满意度需求更大,延迟相同的时间对用户的满意度影响更大,需要作出的补

偿更大,所以在综合考虑能耗成本以及用户满意度之下,其调度幅值较小。

服务器在处理工作负载时产生能耗,因此可以通过负载处理量曲线来反映不同时刻服务器能耗变化情况。数据中心批处理整体到达率 $job1 > job2 > job3 > job4$,但服务器对其各类处理的处理量还要结合不同负载的大小。服务器在不同时刻对各类批处理作业的处理量如图 7 所示。

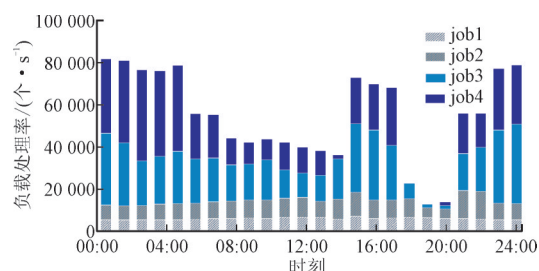


图7 不同时刻批处理负载处理情况

Fig.7 Batch workload processing at different timestamps

由图 7 可知,job3 和 job4 的处理量优化最为明显,受到服务器容量约束之下,利用其时间调度灵活性在电价较低时段大量处理,在电价高峰期时处理量大幅降低,从而优化服务器能耗曲线。

3.2.2 不同负载模型下调度对比分析

为了验证本文构建的批处理负载模型在优化调度中的有效性,以下 3 种模型场景下,将本文模型与主流优化模型进行对比。

场景 1:根据文献[31]的批处理负载模型,采用队列模型,将到达数据中心的批处理负载视为一类,即负载同构,假设其负载可延迟时间为 2 h,负载执行时间为 2 h。

场景 2:批处理负载异构模型,考虑批处理负载之间可延迟时间、执行时间以及到达率等特征的不同以及调度时用户满意度变化。

场景 3:在场景 2 的基础上,考虑不同批处理负载的内部 DAG 结构,充分利用负载内部任务间的优化空间的同时保证执行顺序。

设置 3 种场景下数据中心各设备基本配置相同。场景 1 和场景 3 在不同时刻到达数据中心的负载处理总量相同。场景 2 和场景 3 批处理负载类别以及对应到达率相同。场景 3 为本文所建模型。3 种场景下不同时段批处理负载处理量变化曲线如图 8 所示。

从图 8 中可知,8:00—14:00 以及 18:00—20:00 负载处理量场景 1>场景 2>场景 3,电价较低时段 3 类场景负载处理量大小顺序相反,所以可得场景 3

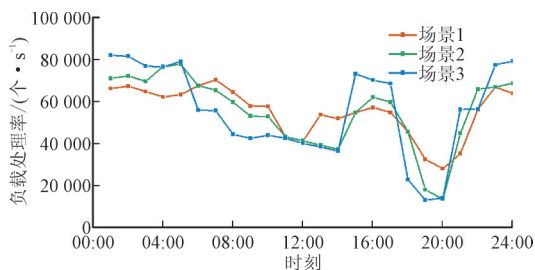


图8 不同场景下服务器批处理负载处理量
Fig.8 Processed batch workloads across different scenarios

的负载模型之下,服务器对批处理负载的处理曲线优化度更高,即服务器的能耗优化更好,这是由于场景3采用了异构工作负载模型,充分考虑了不同批处理负载的可延迟时间、执行时间和到达率的差异,避免了传统同构负载模型的简化假设,使得调度策略能够更精准地匹配负载特性,同时引入了负载内部 DAG 结构的优化,通过分析任务间的依赖关系,充分利用任务执行的时间间隙,进一步挖掘了调度优化的潜力,实现了更细粒度的能耗控制。

3种场景下调度结果如表1所示。

表1 不同负载模型下调度结果对比
Table 1 Scheduling results comparison across different load models

场景	满意度占比/%	综合运行成本/元
1	72.1	54 349.1
2	80.4	52 329.5
3	84.2	49 694.8

3类场景下,场景3综合运行成本相比于场景一降低了8.6%,综合运行成本相比于场景2降低了5.0%,满意度也得到提高,这是由于场景3在负载调度模型中加入用户满意度变化模型,通过动态权衡用户满意度与运行成本,在保证服务质量的前提下,合理分配负载处理时间,避免了因过度延迟导致的用户满意度大幅下降和补偿费用过高。由此说明,本文模型大大提高了运行经济效益,同时也保证了用户的满意度,有利于数据中心的长期发展以及与用户侧的合作。

3.3 日内优化调度分析

日内优化调度后电网的交易量与日前优化调度的对比如图9所示。

由图9可知,日内阶段滚动优化通过协调交互式负载调度以及数据中心内各设备出力,有效减少了因日前预测偏差多导致的日前-日内电网交易量变化。

不同阶段优化运行结果如表2所示。

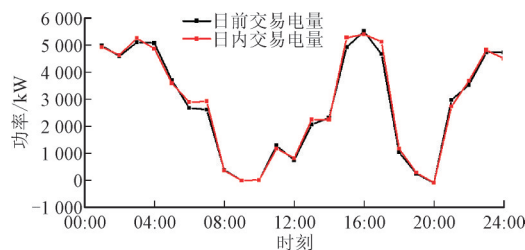


图9 日前-日内数据中心与电网交易电量
Fig.9 Day-ahead and intra-day electricity trading volume between data center and power grid

表2 数据中心两阶段优化运行成本
Table 2 Operational costs of two-stage optimization in data center 元

项目	总运行成本	交互式负载调度补偿	动态服务器开销
日前	49 694.8	0	1 423.2
日内无优化	56 176.1	780.4	1 665.1
日内滚动优化	52 673.2	860.2	1 601.7

因为存在风电光伏以及交互式负载的预测偏差,导致数据中心相比于日前调度计划多负担了13.04%的成本,而通过日内滚动优化可以使得多负担的运行成本下降至5.99%,这是由于在日内滚动优化机制下,通过实时监测交互式负载以及可再生能源出力的变化,动态调整服务器资源分配,避免了因预测偏差导致的资源浪费或不足。同时,异地调度策略充分利用交互式负载空间灵活性,在电价高峰时段将其部分转移至其他数据中心,从而降低了本地服务器的能耗压力,减少高峰时段的用电成本。这些说明本文建立日前-日内两阶段模型,有效降低了预测偏差导致的经济损失。

在日内阶段,由于日内整体交互式负载值高于日前,所以参与调度服务器数量增多,利用日前服务器冗余备份来处理高于日前预测的交互式负载,服务器能耗量也随之增加。同时通过异地调度,保证了交互式负载的即时响应,不会过度抢占服务器资源而导致与批处理负载之间的性能干扰。

4 结论

本文根据数据中心负载特性建立的日前-日内多时间尺度模型,经仿真验证后,得到如下结论:

1) 日前调度中,考虑不同负载基本特征以及内部 DAG 结构,将其扩展到 task 层建立矩阵调度模型,通过对比分析可知,本文建立的模型相比于传统模型调度后经济运行成本降低了8.6%。此外,通过分析不同批处理负载在调度时的转移情况得知,

负载量越大的批处理负载充分利用其内部任务调度优化空间,在调度中显现的效果越好。

2)日内调度中,利用日前服务器冗余备份,同时通过部分交互式负载异地调度以及数据中心内各设备协调出力,有效减少可再生能源出力以及交互式负载数量不确定性带来的经济损失,将因预测偏差导致的额外成本从 13.04% 降低至 5.99%。

未来数据中心将会越来越多,不同数据中心接收的工作负载存在不同,本文建立的模型能够很好地结合当地负载和特征,改变优化周期和时间间隙以应对不同负载情况,在未来将具有更大的发挥空间。

参考文献

- [1] 刘艺.基于负荷转移的数据中心综合能源系统优化研究[D].保定:华北电力大学,2022.
- [2] 祁兵,曹望璋,李彬,等.计及负载特征及响应特性的多数据中心双层优化模型[J].电力系统自动化,2022,46(21):30-41.
QI Bing, CAO Wangzhang, LI Bin, et al. Bi-level optimization model of multipledata centers considering workload features andresponse characteristics[J]. Automation of Electric Power Systems, 2022, 46(21):30-41.
- [3] WANG P, CAO Y J, DING Z H, et al. Stochastic programming for cost optimization in geographically distributed internet data centers [J]. CSEE Journal of Power and Energy Systems, 2022, 8(4):1215-1232.
- [4] 文渐宇,朱继忠,李盛林,等.基于时空协同的多数据中心虚拟电厂低碳经济调度策略[J].电力系统自动化,2024,48(18):56-65.
WEN Xiyu, ZHU Jizhong, LI Shenglin, et al. Low-carbon economic dispatch for multipledata center virtual power plant based on spatio-temporal collaboration [J]. Automation of Electric Power Systems, 2024, 48(18):56-65.
- [5] 龙宇杰,修熙,黄庆,等.基于深度强化学习的电子政务云动态化任务调度方法[J].计算机应用研究,2024,41(6):1797-1802.
LONG Yujie, XIU Xi, HUANG Qing, et al. Scheduling of dynamic dasks in e-governmentclouds using[J]. Application Research of Computers, 2024, 41(6):1797-1802.
- [6] 沈玉明,斯辉,马浩天,等.考虑数据中心和分布式能源接入的配电网双层规划方法[J].全球能源互联网,2023,6(2):116-125.
SHEN Yuming, SI Hui, MA Haotian, et al. A bilevel programming model for distribution networks considering data center and distributed energy access[J]. Journal of Global Energy Interconnection, 2023, 6(2):116-125.
- [7] XU R, MITRA S, RAHMAN J, et al. Pythia: improving datacenter utilization via precise contention prediction for multiple co-located workloads[C]// Proceedings of the 19th International Middleware Conference. New York, USA, 2018:146-160.
- [8] SUN Q H, WU C, LI Z P, et al. Colocation demand response: Joint online mechanisms for individual utility and social welfare maximization[J]. IEEE Journal on Selected Areas in Communications, 2016, 34(12):3978-3992.
- [9] 吕佳伟,张沈习,程浩忠,等.集成数据中心的综合能源系统能量流-数据流协同规划综述及展望[J].中国电机工程学报, 2021, 41(16):5500-5521.
LYU Jiawei, ZHANG Shenxi, CHENG Haozhong, et al. Review and prospect on coordinated planning of energy flow and workload flow in the integrated energy system containing data centers[J]. Proceedings of the CSEE, 2021, 41(16):5500-5521.
- [10] CHEN W, YE K, WANG Y, et al. How does the workload look like in production cloud? Analysis and clustering of workloads on Alibaba cluster trace[C]// Proceedings of the 25th International Conference on Parallel and Distributed Systems (ICPADS 2019). Piscataway, USA, 2019:102-109.
- [11] 张镭,王旭,杨宏坤,等.数据中心集群灵活边界下电力系统分布鲁棒优化调度方法[J].电力系统自动化,2024,48(7):235-247.
ZHANG Ke, WANG Xu, YANG Hongkun, et al. Distributionally robust optimal scheduling method for power system under flexibilityboundaries of data center clusters[J]. Automation of Electric Power Systems, 2024, 48(7):235-247.
- [12] 李彬,杜亚彬,曹望璋,等.考虑风光储互补与工作负载分配的数据中心优化调度[J].现代电力,2022,39(3):356-363.
LI Bin, DU Yabin, CAO Wangzhang, et al. Optimal scheduling of data center considering wind-solar-storage complementary and workload distribution[J]. Modern Electric Power, 2022, 39(3):356-363.
- [13] 崔杨,程禹烽,赵钰婷,等.考虑特性分类批处理负荷可调节能力的数据中心微网灵活性设备分布鲁棒容量配置方法[J].电力自动化设备,2024,44(7):180-188.
CUI Yang, CHENG Yufeng, ZHAO Yuting, et al. Distributionally robust capacity allocation method for flexibility device of data center microgrid considering adjustable capability of characteristic classification batch processing loads[J]. Electric Power Automation Equipment, 2024, 44(7):180-188.
- [14] 陈楷中.数据并行作业负载建模关键技术研究[D].北京:北京工业大学,2022.
- [15] GU Z C, TANG S H, JIANG B L, et al. Characterizing job-task dependency in cloud workloads using graph learning[C]// Proceedings of the 2021 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW). Piscataway, USA, 2021:288-297.
- [16] 王济伟,葛浙奉,蒋从锋,等.混部数据中心负载特征及其任务调度优化分析[J].计算机工程与科学,2020,42(1):8-17.
WANG Jiwei, GE Zhefeng, JIANG Congfeng, et al. Workload characterization and task scheduling optimization of co-located internet data centers[J]. Computer Engineering & Science, 2020, 42(1):8-17.
- [17] 曹雨洁,丁肇豪,王鹏,等.能源互联网背景下数据中心与电力系统协同优化(二):机遇与挑战[J].中国电机工程学报,2022,42

- (10):3512-3527.
- CAO Yujie, DING Zhaohao, WANG Peng, et al. Coordinated operation for data center and power system in the context of energy internet (II): Opportunities and challenges[J]. Proceedings of the CSEE, 2022, 42(10): 3512-3527.
- [18] 李文信, 齐恒, 徐仁海, 等. 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(4): 600-617.
- LI Wenxin, QI Heng, XU Renhai, et al. Data center network flow scheduling progress and trends[J]. Chinese Journal of Computers, 2020, 43(4): 600-617.
- [19] JIANG C F, HAN G J, LIN J B, et al. Characteristics of co-allocated online services and batch jobs in internet data centers: A case study from Alibaba Cloud[J]. IEEE Access, 2019, 7: 22495-22508.
- [20] 刘迪, 曹军威, 刘明爽. 分布式数据中心信息能量协同优化策略[J]. 清华大学学报(自然科学版), 2022, 62(12): 1864-1874.
- LIU Di, CAO Junwei, LIU Mingshuang. Collaborative optimization strategy of information and energy for distributed data centers[J]. Journal of Tsinghua University (Science and Technology), 2022, 62(12): 1864-1874.
- [21] 易文飞, 朱卫平, 郑明忠. 计及数据中心和风电不确定性的微电网经济调度[J]. 中国电力, 2024, 57(2): 19-26.
- YI Wenfei, ZHU Weiping, ZHENG Mingzhong. Economic dispatch of microgrid considering data center and wind power uncertainty[J]. Electric Power, 2024, 57(2): 19-26.
- [22] 刘祎泽, 向月. 计及负荷转移需求响应的低碳数据中心光储容量优化配置[J]. 电力自动化设备, 2024, 44(7): 149-155.
- LIU Yize, XIANG Yue. Optimal allocation of photovoltaic and storage capacity in low carbon data center considering load transfer demand response[J]. Electric Power Automation Equipment, 2024, 44(7): 149-155.
- [23] 苏娟, 董彦君, 赵晶, 等. “东数西算”背景下多数据中心联合消纳可再生能源途径研究综述[J]. 高电压技术, 2024, 50(1): 55-64.
- SU Juan, DONG Yanjun, ZHAO Jing, et al. Research of the approach to consuming renewable energy across multiple data centers in the context of “East Data, West Calculation” [J]. High Voltage Engineering, 2024, 50(1): 55-64.
- [24] 魏震波, 张海涛, 魏平桢, 等. 考虑动态激励型需求响应的微电网两阶段优化调度[J]. 电力系统保护与控制, 2021, 49(19): 1-10.
- WEI Zhenbo, ZHANG Haitao, WEI Pingan, et al. Two-stage optimal dispatching for microgrid considering dynamic incentive-based demand response[J]. Power System Protection and Control, 2021, 49(19): 1-10.
- [25] 曹望璋, 李彬, 祁兵, 等. 计算资源共享模式下的托管式数据中心群两阶段鲁棒优化模型[J]. 电网技术, 2022, 46(10): 4102-4115.
- CAO Wangzhang, LI Bin, QI Bing, et al. Two stage robust optimization model of colocation data centers cluster under computing resource sharing mode[J]. Power System Technology, 2022, 46(10): 4102-4115.
- [26] 唐伟, 单葆国, 郑海峰, 等. 数据中心两阶段源荷协同优化调度研究[J]. 电力系统及其自动化学报, 2023, 35(4): 49-58.
- TANG Wei, SHAN Baoguo, ZHENG Haifeng, et al. Research on two-stage source-load coordinated optimal scheduling of data centers[J]. Proceedings of the CSU-EPSA, 2023, 35(4): 49-58.
- [27] YU J Y, TONG W, LV P Z, et al. TERMS: task management policies to achieve high performance for mixed workloads using surplus resources[J]. Journal of Parallel and Distributed Computing, 2022, 170: 74-85.
- [28] 王康瑾, 贾统, 李影. 在离线混部作业调度与资源管理技术研究综述[J]. 软件学报, 2020, 31(10): 3100-3119.
- WANG Kangjin, JIA Tong, LI Ying, et al. State-of-the-art survey of scheduling and resource management technology for colocation jobs[J]. Journal of Software, 2020, 31(10): 3100-3119.
- [29] JIANG C F, QIU Y T, SHI W S, et al. Characterizing co-located workloads in Alibaba cloud datacenters[J]. IEEE Transactions on Cloud Computing, 2022, 10(4): 2381-2397.
- [30] 杨秀, 张相寅, 黄海涛, 等. 基于多智能体近端策略网络的数据中心双层优化调度[J]. 南方电网技术, 2025, 19(4): 107-121+131.
- YANG Xiu, ZHANG Xiangyin, HUANG Haitao, et al. Double layer optimization scheduling of data center based on multi-agent proximal strategy network[J]. Southern Power System Technology, 2025, 19(4): 107-121+131.
- [31] DING Z H, CAO Y J, XIE L Y, et al. Integrated stochastic energy management for datacenter microgrid considering waste heat recovery[J]. IEEE Transactions on Industry Applications, 2019, 55(3): 2198-2207.

(责任编辑 钱文姝)

收稿日期: 2024-11-14; 修改日期: 2025-05-26

作者简介: 帅文利(2000), 女, 四川眉山人, 硕士研究生, 主要研究方向为数据中心优化调度。E-mail: 2313203936@qq.com

赵晓莉(1979), 女, 通信作者, 安徽淮南人, 讲师, 主要研究方向为电力系统规划与稳定控制、新能源接入电力系统的运行分析。E-mail: zhaoxiaoli@shiep.edu.cn