Electrical Power and Energy Systems 159 (2024) 110032

Contents lists available at ScienceDirect

International Journal of Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Energy-aware coordinated operation strategy of geographically distributed data centers

Shibo Zhou, Ming Zhou, Zhaoyuan Wu*, Yuyang Wang, Gengyin Li

School of Electrical and Electronic Engineering, North China Electric Power University, Changping District, Beijing 102206, China

A R T I C L E I N F O

Keywords:
Coordinated cooperation
Geographically distributed internet data centers
Multiple coupled scheduling methods
Profit-sharing mechanism
Spatial–temporal flexibility

A B S T R A C T

As digital technology continues to advance rapidly, the substantial operational electricity costs associated with the extensive deployment of globally distributed internet data centers have become a significant concern. Addressing this challenging issue is crucial for the future expansion of internet data centers. In this paper, a coordinated strategy is proposed that leverages the spatial–temporal flexibility of internet data centers and their interaction with the power system. Firstly, an energy-aware internet data center loading model is introduced that incorporates multiple coupled scheduling methods to maximize the utilization of internet data centers' flexibility potential. Secondly, the proposed internet data centers loading model is integrated into a coordinated optimization model aimed at achieving higher profits through cost reduction in response to varying electricity prices across multiple regions. To facilitate the coordinated operation of geo-graphically internet data centers, an incentive profit-sharing mechanism based on Nash Bargaining theory is designed to fairly distribute coordination profits based on the contribution ratio of each data center in coordination. Through comprehensive case studies utilizing real-world datasets, the results demonstrate that the internet data center loading model, considering multiple coupled scheduling methods, can enhance internet data centers' flexibility potential, resulting in a significant 11.73% reduction in total cost, with the coupling relationship accounting for 5.05% of this reduction. Furthermore, the proposed profit-sharing mechanism effectively fosters coordination and enthusiasm among geographically distributed internet data centers.

1. Introduction

1.1. Background

With the rapid development of global digital technology and the devastating impact of the coronavirus disease since 2019, the digital economy has emerged as a crucial catalyst for countries to achieve economic recovery and drive transformation and upgrading [1]. The number of servers newly put into use in internet data centers (IDCs) has grown steadily. However, being an energy-intensive industry, the considerable power consumption costs pose a significant challenge to its expansion. In 2020, data centers were responsible for 2.7 % of the European Union's total electricity demand. This figure is projected to increase to 3.2 % by 2030 [2].Thus, resolving the cost issue associated with the substantial power consumption in IDCs has become the primary focus of current research endeavors.

With the deregulation of the electricity market, the prices and incentives have changed the traditional operation strategy of geo-distributed IDCs. They no longer focus on pursuing lower power consumption through energy management methods. Instead, they achieve more economically viable power consumption strategies by flexibly scheduling their loads to cater for the differences in regional electricity prices [3]. Especially with the innovation and development of optical fiber and network architectures, an IDC not only can effectively transfer its workloads to other time slots locally, where electricity prices are relatively low by utilizing the batch workload scheduling (BWS) method [4], but also can transfer its workloads to other geo-distributed IDCs where electricity supply are relatively abundant by utilizing the geographical workload balancing (GWB) [5] for the minimum cost expenditure. For example, China recently proposed the strategy of " Channel Computing Resources from The East to The West ": guiding the rich data resources in the east to the west with rich natural resources helps solve the dilemma of unbalanced computing resources and electricity resources between the east and the west as well as realizes IDCs' operational cost optimization by more renewable energy accommodation. In addition, some studies focus on the management of cooling system to deal with its huge power consumption in an IDC, the thermal

* Corresponding author at: School of Electrical and Electronic Engineering, North China Electric Power University, Changping District, Beijing 102206, China.
E-mail address: wuzy@ncepu.edu.cn (Z. Wu).

https://doi.org/10.1016/j.ijepes.2024.110032
Received 5 September 2023; Received in revised form 11 March 2024; Accepted 7 May 2024
Available online 13 May 2024
0142-0615/

Nomenclature

A. Sets

N^I	Number of IDCs
N^w	Number of wind power units
N^h	Number of thermal power units
N^l	Number of system load except the IDCs
T	Set of time slots
K	Set of thermal power unit output stages

B. Symbols

λ	Delay-sensitive workload
δ	Coefficient of performance of air conditioning
τ	Delay tolerance of delay-sensitive workload
ρ	Air density
ξ	Profits distribution result
φ	Cost coefficient of wind
θ	Cost coefficient of workload migration
κ	Coefficient of temperature attenuation
e	Rated power consumption per server
m	Servers in an internet data center
R	Profits in the coordination
u	Average service rate of servers
χ	Delay-tolerant workload
υ	Delay tolerance of delay-sensitive workload
β	Coefficient of performance of air conditioning
η	Contribution ratio of the coordination
ϖ	Coefficient of aggregator fee
γ	Locational marginal price
ε	Carrying heat per unit air
C	Cost
f	Flow rate per unit air
P	Power consumption
Tem	Temperature
x	Operation state

C. Superscript

cool	Cooling system
th	Thermal power unit
in	Indoor environment
oth	Other components of an internet data center
wi	Wind power unit
out	Outdoor environment
pro	Power procurement
sur	Cooperative surplus
fac	Cooling facilities
idle	Servers at an idle state
load	System load except for the IDCs
wor	Workload processing in the minimum QoS
mig	Workload migration
peak	Servers at a peak state
re	Redundant system

D. Subscript

α_i	Index of coefficient of performance of air conditioning
i	Index of internet data centers
t	Index of time slots
a_i	Index of thermal power cost coefficient
k	Index of stages of thermal power unit

E. Abbreviations

BWS	Batch workload scheduling
GWB	Index of internet data centers
PUE	Power usage effectiveness
QoS	Quality of service
CSP	Cloud service provider
IDC	Internet data center
LMP	Locational marginal price
TIAC	Thermal inertia of computer room air conditioning

inertia of computer room air conditioning (TIAC) is proposed to schedule the cooling plant outputs by collecting the waste cold released by overcooled air and raised metal floor storing thermal masses from working servers in two adjacent time slots for maximum savings [6,7]. Therefore, large-scale IDCs with spatial–temporal flexibility can optimize operation efficiency and effectively reduce the power procurement costs, when these GWB, BWS, or TIAC scheduling methods are properly implemented [8].

Moreover, the independent IDC cannot directly coordinate with other IDCs for information interchange due to the high workload privacy, so the IDCs often choose the operation mode of 'third-party participation' to be a platform for coordination [9]. Specifically, geo-distributed IDCs submit their information about workload arriving to a third-party company called cloud service provider (CSP) that plays a role of information platform for the optimal operation state, then the CSP, as a trading platform, conducts power trading with the power system. This operation mode ensures that the IDC and the power system can make decisions to maximize their own interests. Only necessary information is exchanged, protecting the privacy of each member in the coordination with low-risk costs. However, when geo-distributed IDCs participating in coordination to seek the maximization of collective interests on cost reduction based on varying electricity prices across multiple regions, the interest of one IDC in coordination may be infringed [10]. The reason for this contradictory phenomenon is that large workloads of other IDCs are transferred to this IDC, leading to a higher power procurement cost than that before participating in coordination. This will frustrate the enthusiasm of some members to participate in the coordination, so it is necessary to design a mechanism of profits sharing within the coordination to achieve a win–win situation of the whole and part.

1.2. Literature survey

Recent extensive interests in IDCs have been partially intrigued by expectations of the coming digital future, but the enormous electricity costs hinder the large-scale construction of IDCs which raise the prospect of IDCs participating in the deregulated electricity market interaction to solve this problem. Many studies have focused on building the IDC loading model for power consumption, which helps to describe IDC electric features for better scheduling. Some of the studies in this area focus on (1) measuring the power-supply voltage and frequency of servers to describe internet technology (IT) equipment power consumption [11,12]; (2) leveraging the indicator power usage efficiency (PUE) to describe power consumption outside of IT equipment [13,14]. However, the literature [15] discusses the availability of idle servers in redundant systems, despite improvements in server performance and workload prediction accuracy. Moreover, the literature [6] shows that the indoor heat heating exhibits a hysteresis effect due to the law of thermal inertia. Therefore, cooling facilities do not need to operate at a constant power level. These factors are all overlooked in studies about the IDC loading model.

Moreover, the purpose of building IDC electric loading model is to better make economic and effective operation strategy. Many studies have made full use of the spatiotemporal flexibility of IDCs for energy saving and cost reduction. Firstly, Integrated energy system is usually considered to exploit heterogenous energy resources for minimizing power consumption for IDCs [16,17]. The literature [18] proposes an optimal sizing method for the energy station in the multi-energy system integrated with data centers to reduce the power consumption. Secondly, the diversity of electricity prices in deregulated electricity market is utilized by IDCs for arbitrage in curtailing power costs [19,20]. The literature [21] applies real-life data traces of the electricity price, renewable energies and heating demand for total cost reduction and energy usage efficiency in the whole operating system. Thirdly, the abundant and cheap renewable energy also provides a novel opportunity for data centers to save energy and reduce costs [22,23]. The literature [24] suggests that the usage of renewable energy resources with appropriate server selection and consolidation can mitigate the energy related issues in cloud environment. While these studies are inspiring, they often use energy-aware scheduling methods for IT equipment. Few literatures have focused on the relationship between the flexibility of IT equipment and the cooling system, particularly when the effectiveness of the TIAC is affected by the implementation of the GWB and the BWS.

In addition, coordination between geo-distributed IDCs to take advantage of local resources, low electricity prices, and relevant favorable policies has become a major trend today when encountering different resource endowments in multiple regions. The literature [25] proposes an aggregation method for proliferated small-size IDCs to apply spatial load regulation potentials based on the diversity in regional electricity prices. The literature [26] proposes a pricing scheme tailored for geo-distributed green data centers, from a multi-leader (smart grids) single-follower (cloud) game point of view, to get the economic profits of the proposed real-time pricing. Coordination inevitably takes place in a way that involves the sharing of profits between members. The literature [27] demonstrates that each data center aggregator should compensate its customers according to the established agreement to ensure active participation of IDCs in demand response events. These studies do not take into account the coordination's profit-sharing or the varying marginal contribution ratios of the coordinated parties to achieve a fair distribution of results. This may lead to incentive incompatibility, where the interests of some members are sacrificed to maximize collective interests.

1.3. Contributions

IDCs are significant load resources that offer substantial potential for electricity market participation due to their operational characteristics. They allow for task deferral in time and spatial shifting, making them highly flexible. However, current studies have several short-comings. Firstly, they overlook the cooling system's flexibility potential by solely relying on PUE to measure energy consumption, neglecting heat inertia. Secondly, there is a shortage in modelling for coordinated interaction between IDCs and the grid, as well as underutilization of IDCs' spatiotemporal flexibility in response to electricity prices. Thirdly, while IDCs' spatial flexibility relies on coordination among geo-distributed IDC, research often overlooks prof-it-sharing among internal members, risking incentive incompatibility. The above discussions motivate us to answer the following key questions in this paper.

1) How to portray the spatiotemporal flexibility of the IDC and their impact on participation in market transactions.
2) How can the mutually independent game interaction between IDCs and power grids be reflected in market transactions?
3) How to design a rational profit-sharing mechanism to avoid the effects of potential incentive incompatibility and promote coordination motivation?

To address the aforementioned questions, this paper proposes a coordinated optimization model for exploiting the geo-distributed IDCs' spatial–temporal flexibility potentials by three coupled scheduling methods for coordination profits in power transactions. Then, a profit-sharing mechanism is designed to share these profits according to each IDC's contribution ratio in the coordination. The contributions are threefold:

1) A novel energy-aware IDC loading model is established that considers the inertia effect of ambient temperature, and incorporates three scheduling methods (BWS, GWB and TIAC) to quantify the power consumption of various components, which departs from the traditional use of Power Usage Effectiveness (PUE) as the only metric. Additionally, considering advancements in server performance and workload prediction accuracy, the scheduling of idle servers within redundant systems is addressed to ensure optimal server utilization, thereby achieving higher Quality of Service (QoS) within an IDC.
2) A coordinated optimization model among the geo-distributed IDCs utilizing three energy-aware scheduling methods is proposed, the CSP being a third-party platform, and the power system with multiple dynamically changing electricity prices. Compared to existing research, this operational mode ensures that the IDC and power system can maximize their own interests through an interactive game, where the IDC is no longer a 'price taker' but a 'price maker'. Only necessary information is exchanged, protecting the privacy of each member in the coordination with low-risk costs.
3) An incentive profit-sharing mechanism based on Nash Bargaining theory is designed to facilitate the coordination of geographically distributed data centers. The proposed mechanism compensates for the extra power procurement costs caused by workload migration, and shares residual profits among geo-distributed IDCs based on their contribution ratios. Compared to other profit-sharing methods, the analytical solution of the Nash Bargaining model can be applied to larger-scale systems, and the results of profit-sharing also follow the principle of 'you do the most, you get the most'. Therefore, it is verified that the profit-sharing mechanism is reasonable and fair.

2. Problem description

2.1. Spatial-Temporal flexibility of the IDC

Depending on the characteristic of its workloads, IDC has the spatial–temporal flexibility potential for scheduling its power demand across regions and time slots, which contributes to cost reduction by the arbitrage among different electricity prices. The details are shown in Fig. 1.

1) Temporal flexibility.

Some workloads have a high delay tolerance, with a maximum response time of a few minutes or a few days, which named delay-tolerant workloads [24], such as image processing, large-scale data analysis, etc. Temporal flexibility refers to the delay-tolerant workload operation that can be deferred within a certain time. Moreover, due to the influence of thermal inertia, cooling facilities does not need to maintain constant power for operation, but dynamically adjusts its output in the form of thermal gradient in two adjacent time slots, which can also be referred to temporal flexibility [28].

2) Spatial flexibility.

With development of optical fiber and network architecture, some workloads, such as live video, data query and business transaction application, have a maximum response time of a few milliseconds, which named delay-sensitive workloads [24]. Spatial flexibility refers to the delay-sensitive workloads that can be processed in any severs of any IDC, regardless of geographic region, as long as it satisfies the requirement of minimum Quality of service (QoS) which is considered into the

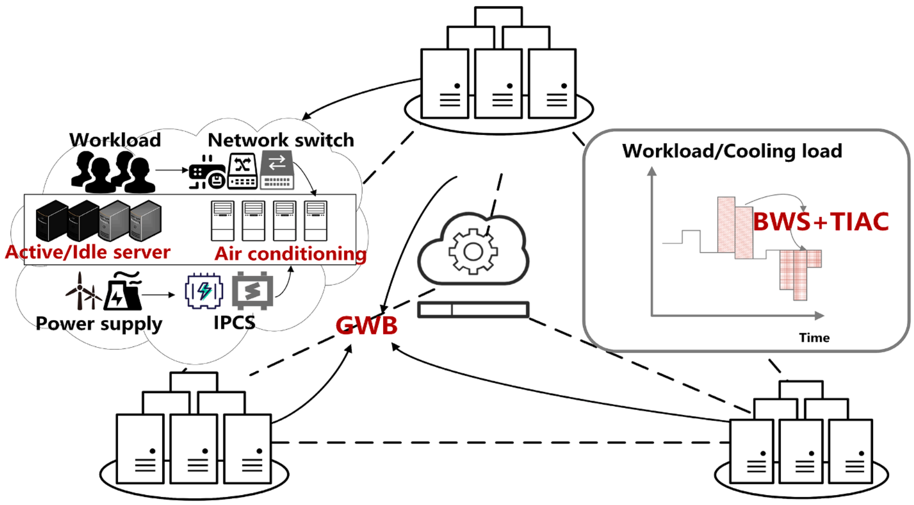

Fig. 1. Spatial-temporal flexibility of geo-distributed IDCs.

IDC loading model [5].

2.2. Coordination framework

To exploit geo-distributed IDCs' full flexibility potentials and incentivize coordinated operation, this paper proposes a coordination framework among geo-distributed IDCs, which can be divided into the following two parts, i.e., coordination operation and profit-sharing. The proposed coordinated operation framework is shown in Fig. 2, which is discussed as follows.

1) The coordinated optimization model is the first part. In this part, a bi-level optimization model is designed to achieve more coordination profits. The upper level is controlled by the geo-distributed IDCs for activating spatial–temporal flexibility potentials, and the lower level is guided by the power system. The upper level achieves the optimal operation costs through spatial–temporal flexibility with the help of the CSP as an information flatform, and then the power demand is submitted to the CSP. The lower level receives the power demand from the CSP and then calculates the locational marginal prices (LMPs). Finally, the lower-level feedbacks these price signals to the geo-distributed IDCs as the guidance for the strategies of workload migration with the help of the CSP as a trading platform.
2) The reasonable profit-sharing mechanism is the second part. In this part, a reasonable profit-sharing mechanism is proposed to avoid the possibility of incentive incompatibility, which firstly compensate the possible loss of participation in coordination and then distribute the coordination profits according to the contribution ratio of each IDC after paying a certain amount of fees to the CSP. Moreover, the Nash Bargaining model's solution is analytical, making it possible to implement this method in larger-scale systems.

3. Energy-aware IDC loading model

In this section, a novel energy-aware IDC loading model is established that incorporates the inertia effect of ambient temperature and considers three scheduling methods to quantify the power consumption of various components. First, the workloads are classified into delay-tolerant and delay-sensitive workloads according to their delay tolerance. Next, this paper defines an hour as a time slot for updating decisions related to cooling system control, workload distribution, and power system scheduling [10]. Three energy-aware scheduling methods

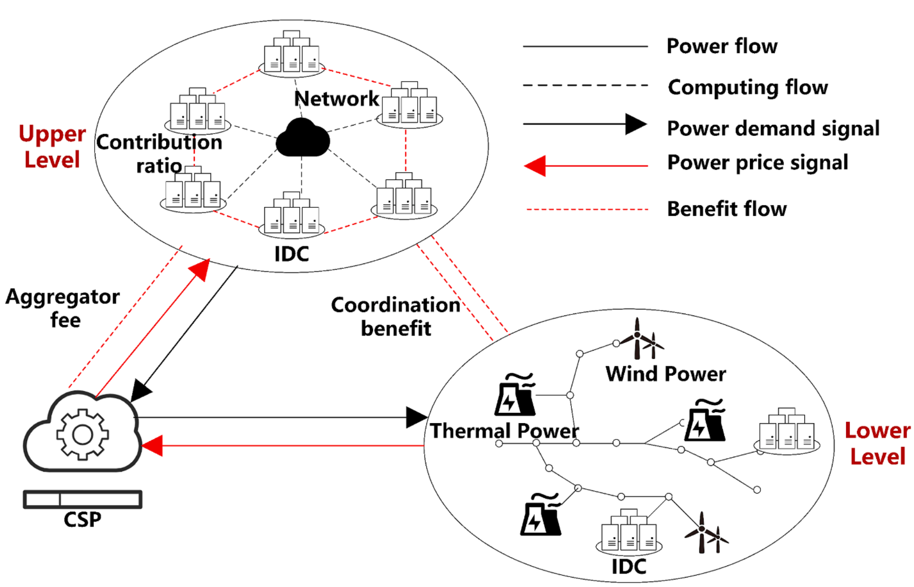

Fig 2. Coordination framework among geo-distributed IDCs.

(GWB, BWS, TIAC) are considered to regulate the power consumption of an IDC. Finally, an energy-aware IDC loading model is proposed.

Fig. 3 illustrates an epitome of the components and power-heat flow in an IDC [29]. Firstly, Internal power conditioning system and power distribution unit as parts of infrastructure are aimed to ensure that power flow is supplied to the servers. Then, servers in IT equipment and redundant system execute calculation, storage and redundance services for the workloads with huge power consumption. Finally, the cooling facilities provided by power from grid can schedule their outputs to cool the heat released by servers to maintain a reasonable temperature range.

As shown in Fig. 3, all components consume the electric power from power grid. The power consumption of each IDC can be characterized by four parts, including workload processed in IT system, redundant system, cooling system and other auxiliary facilities. The power consumption loading model of IDC i is formulated in (1).

$$P_{it}^{IDC} = P_{it}^{wor} + P_{it}^{re} + P_{it}^{cool} + P_{it}^{oth} \tag{1a}$$

3.1. IT equipment and redundant system

Based on the above the components in an IDC, this research on IDC loading model will not only considers power consumption of servers, but also considers the power consumption of the infrastructure such as an internal power conditioning system and distribution unit and so on. Moreover, When the workloads are migrated across regions or different time slots, workloads may be concentrated to a time slot or an IDC. It may aggravate the problem that the queueing is too long to satisfy the delay requirement. Therefore, it is necessary to utilize the servers affiliated in the redundant system which is usually idle due to improvement of the quality of servers [15]. Accurately expressing the power features of an IT system is crucial as it constitutes the most significant part of its power consumption. The analytic expression is shown as follows:

$$P_{it}^{wor} = (e_i^{inf} + e_i^{idle})m_{it}^{wor} + (e_i^{peak} - e_i^{idle})(\frac{\lambda_{it} + \chi_{it}}{u_i}) \tag{1b}$$

$$\begin{aligned} P_{it}^{re} = {} & (e_i^{inf} + e_i^{idle})m_{it}^{re} + (e_i^{peak} - e_i^{idle})(m_{it}^{peak} - \frac{\lambda_{it}}{u_i(m_{it}^{ws} + m_{it}^{rs})}) \\ & + (e_i^{peak} - e_i^{idle})(m_{it}^{peak} - \frac{\chi_{it}}{u_i(m_{it}^{wt} + m_{it}^{rt})}) \end{aligned} \tag{1c}$$

$$m_{it}^{ws} = \lambda_{it}/(u_i - 1/\nu) \tag{1d}$$

$$m_{it}^{wt} = \chi_{it}/u_i \tag{1e}$$

$$m^{ws} + m^{wt} = m^{wor} \tag{1f}$$

$$m^{rs} + m^{rt} = m^{re} \tag{1g}$$

$$m_{it}^{peak} \leqslant m_{it}^{wor} + m_{it}^{re} \tag{1h}$$

Based on the QoS constraint and maximum server utilization, constraints (1b) - (1c) describe the power consumption of IT system, where the workloads are processed before the delay deadline. Constraint (1b) is a typical regression model showing the correlation between the electricity consumption and the computing utilization [29]. In constraint (1c), the first part is the idle electricity consumption for redundant

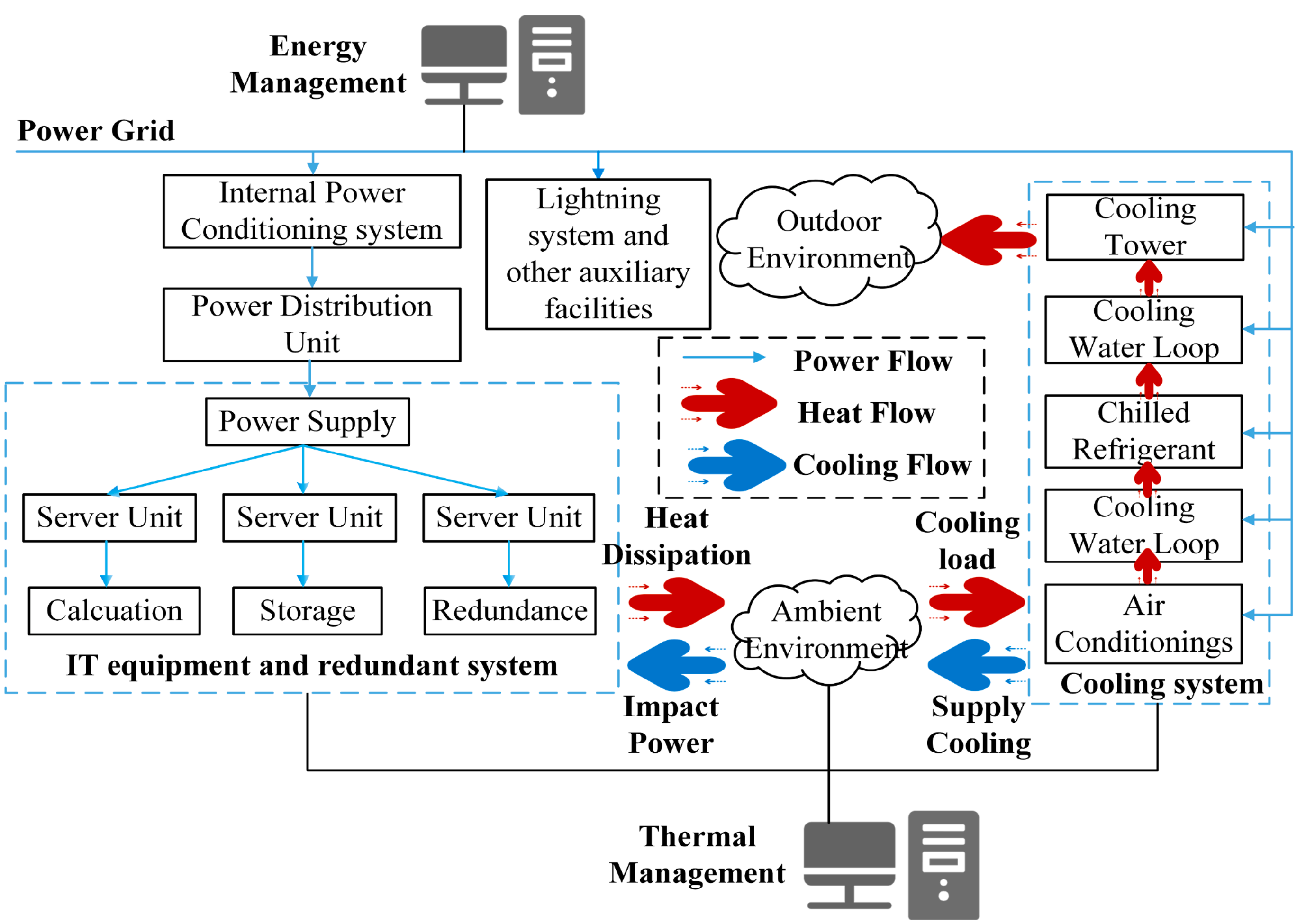

Fig. 3. Components and power-heat flow in an IDC.

servers, while the second part and the third part are the extra electricity consumption which utilize redundancy in computing resources for higher efficiency. Constraint (1d) describes the power relationship between active servers and processed delay-sensitive workloads in a M/M/1 queue model that represents the response time [30]. Constraint (1e) describes the power relationship between active servers and processed delay-tolerant workload. Constraints (1f) - (1 h) describe the relationship among different types of servers.

3.2. Cooling load

In previous studies, the power consumption of cooling load is presented in form of PUE which means that cooling facility outputs are not controlled by default. However, the heat has a hysteresis effect in practice, so the outputs of cooling facility can be scheduled in different time slots. The power consumption of cooling system is decided by cooling facilities which is shown as follows:

$$P_{it}^{cool} = \beta_i P_{it}^{fac} + P_{\min}^{cool} \tag{1i}$$

$$\beta_i = \frac{1}{\alpha_1 \left(Tem_{it}^{Out}\right)^2 + \alpha_2 Tem_{it}^{Out} + \alpha_3} \tag{1j}$$

$$P_{\min}^{cool} \leqslant P_{it}^{cool} \leqslant P_{\max}^{cool} \tag{1k}$$

where constraints (1i) describe the relationship between the power consumption of the cooling system and the cooling facility outputs. Constraint (1j) describes the coefficient of performance of air conditioning [31]. In addition, the heat tolerance of the servers is limited, so constraint (1k) is needed to limit the indoor temperature.

3.3. Scheduling methods

The spatial–temporal flexibility potential of IDCs requires specific scheduling methods to show its advantages. This model implements three scheduling methods uniformly to maximize flexibility in the IDC. The specific expression is shown as follows:

$$\sum_{i=1}^{N^I} \lambda_{it} = \sum_{i=1}^{N^I} \lambda_{it}^{ina} \tag{1l}$$

$$\sum_{i=1}^{N^I} \sum_{t=1}^{t} \chi_{it} \leqslant \sum_{i=1}^{N^I} \sum_{t=1}^{t} \chi_{it}^{ina} \forall t \in T \tag{1m}$$

$$\sum_{i=1}^{N^I} \sum_{t=1}^{t+\tau} \chi_{it} \geqslant \sum_{i=1}^{N^I} \sum_{t=1}^{t} \chi_{it}^{ina} \forall t \in [1, T-\tau] \tag{1n}$$

$$Tem_{it}^{In} - Tem_{i(t-1)}^{In} = \frac{-P_{it}^{fac} + P_{it}^{wor} + P_{it}^{re} + P_{it}^{oth}}{\rho \varepsilon f_i} + Tem_{it}^{Out}\left(1 - e^{\kappa}\right) \tag{1o}$$

$$Tem_{\min}^{in} \leqslant Tem_{it}^{in} \leqslant Tem_{\max}^{in} \tag{1p}$$

where constraint (1l) indicates that workloads remain unchanged before and after spatial migration, which is associated with GWB. Constraints (1 m) - (1n) indicates that workloads can be migrated in its maximum delay tolerance, which associated with BWS. Constraints (1o) - (1p) describe the thermal inertia effect, which is associated with TIAC. It is essentially an extension of the law of conservation of energy in thermodynamics [28]. It demonstrates that the effectiveness of TIAC is associated with the power consumption of IT equipment (P_{it}^{wor} and P_{it}^{re}), which is affected by GWB and BWS.

In comparison to the traditional IDC loading model, the proposed model, represented by (1), has three main advantages: 1) The impact of ambient temperature on cooling system power consumption is analyzed, while other forms of power consumption, apart from the IT system, are not described simply with an indicator such as PUE, for greater accuracy; 2) With improved server quality and workload prediction, scheduling methods can now utilize more idle resources, extending the scheduling range, including the use of redundant systems; 3) Three coupled scheduling methods are fully taken into account to maximize the spatial–temporal flexibility potentials of geo-distributed IDCs.

4. The coordinated optimization model

To represent the coordination operation among geo-distributed IDCs based on varying electricity prices, the information interaction is essential. In this model, a CSP is chosen as the ‘third-party’ entity to help the coordination among geo-distributed IDCs. This model implicitly assumes a power market with full competition, in which the power system can dynamically give the LMPs according to the output of marginal units, and the IDCs can accurately present their power demand after workload migration by energy-aware scheduling methods.

4.1. The Upper-Level model

The upper-level model is the coordinated operation model for multiple geo-distributed IDC, formalized by (2):

$$\min C^{pro} + C^{mig} \tag{2a}$$

$$C^{pro} = \sum_{t \in T} \sum_{i \in N^I} P_{it}^{IDC} \gamma_{it} \tag{2b}$$

$$C^{mig} = \frac{1}{2} \theta^{mig} \sum_{i \in N^I} \left| \sum_{t \in T} P_{it}^{wor} - \sum_{t \in T} P_{it}^{nwor} \right| \tag{2c}$$

Subject to:

Constraints (1a)– (1p) (2d).

The upper-level objective function (2a), which includes the extra cost of workload migration among geo-distributed IDCs, minimizes the IDC operation costs. γ_{it} represents the LMP where IDC i is located. γ_{it} is provided by the lower model. Constraint (2c) represents the price of workload migration which is the electricity consumption of core switching equipment in backbone network unit workload migration [32]. P_{it}^{nwor} presents the electricity consumption of IT equipment without coordination. Constraints (2d) describes the energy-aware IDC loading model for considering the three scheduling methods.

4.2. The Lower-Level model

The lower-level model is the power system operation model, which aims to minimize operating costs, as shown in (3).

$$\min \sum_{t=1}^{T} \left(\sum_{th}^{N^h} C_t^{th} + \sum_{wi}^{N^w} C_t^{wi} \right) \tag{3a}$$

$$C_t^{th} = a_1 \left(P_t^{th}\right)^2 + a_2 P_t^{th} + a_3 \tag{3b}$$

$$C_t^{wi} = \varphi^{wi} P_t^{wi} \tag{3c}$$

Subject to:

$$P_t^{th} = P_{\min}^{th} x_t^{th} + \sum_{k=1}^{K} P_{kt}^{th} \tag{3d}$$

$$0 \leqslant P_{kt}^{th} \leqslant P_{\max}^{th} \tag{3e}$$

$$\sum_{w \in N^w} P_t^{wi} + \sum_{h \in N^h} P_t^{th} = \sum_{i \in N^I} P_{it}^{IDC} + \sum_{l \in N^l} P_t^{load} [\gamma_{it}] \tag{3f}$$

The whole cost of the lower-level objective function (3a) includes the combined operation cost of thermal power and wind power. Constraints

(3d)–(3e) represent the thermal power generating unit and wind power turbine constraints. P_{it}^{IDC} in constraint (3f) represents the power consumption of IDC i, which is provided by the upper-level model. γ_{it} presents the dual variable of constraint (3f), and the derivation of it is shown in **APPENDIX**. However, some traditional constraints of thermal power units will not be described for space reasons [33].

In a word, the set of formulas (2) and set of formulas (3) form this bi-level model to realize the optimum of both parties. In addition, for the nonlinear expression in the model, such as the absolute value in Eq. (2c), the method of 0–1 variable is used to transform it into a linear constraint for solving. The derivation is shown in **APPENDIX.** Thus, the above model can be transformed into a mixed integer linear programming model, and the specific solution process is shown in **Section III-C.**

4.3. Solution method

In the bi-level model proposed in this paper, both the upper-level and lower-level models are mixed integer linear programming models, which can be solved iteratively. The flowchart of the solution process is shown in Fig. 4, and can be presented as the following steps:

Step1: Set the basic parameters of power system, including wind power forecast, operation parameters of thermal units, electricity load forecast, network topology and so on.

Step2: Input the initial parameters of geo-distributed IDCs, including the number of arriving workloads, the maximum delay of the delay-tolerant, the outside temperature and so on.

Step3: Initialize the number of iterations, $G = 0$.

Step4: The upper-level model for the operation of IDCs is solved with the goal of minimizing the total cost of geo-distributed IDCs, and the electric power consumption P_{it}^{IDC} is transmitted to the lower-level model.

Step5: The lower-level model for the power system is solved with the goal minimizing the total cost of the operation of power system, and achieve the locational marginal price γ_{it} according to the duality theory.

Step6: Determine whether the difference between the total cost of geo-distributed IDCs in **iteration G** and **iteration G + 1** is less than the convergence criterion ΔΓ (C(G + 1)-C(G) < 0.001). If yes, terminate the iterative solution and output the operation decision information of IDCs and locational marginal prices of the power system; if no, the locational marginal prices are fed back to the upper-level model and goes to **Step 3** to continue the next iteration until the convergence criterion is met, then output the results.

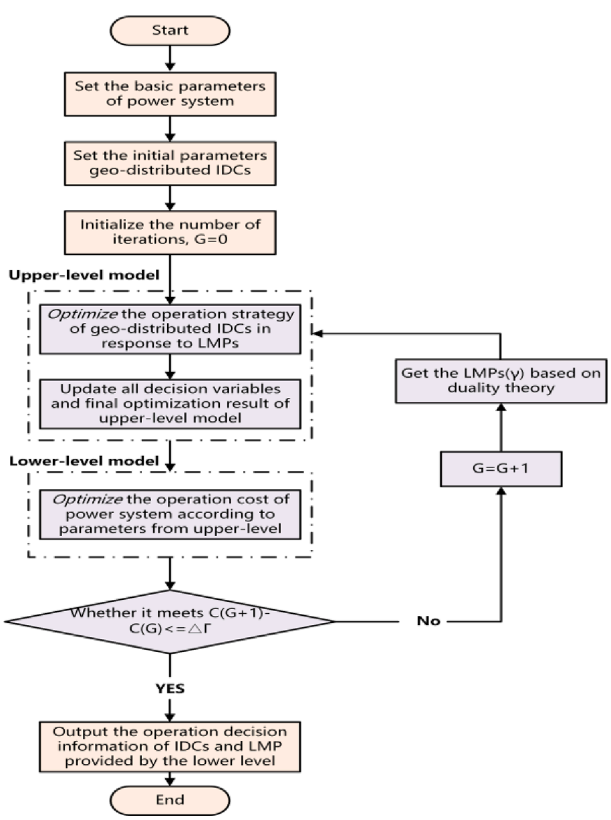

Fig. 4. Flowchart of the solution of the bi-level model.

5. Profit-sharing mechanism

Now, geo-distributed IDCs tend to make a coordination to maximize the utilization of local advantages by spatial–temporal flexibility. A fair and reasonable coordination should guarantee the interest of each member depending on its contribution. This section proposes the contribution ratio as a means of identifying the relevant contributions of different IDCs in a coordination. On the ground of these indices, an asymmetric Nash bargaining model is formulated for profit-sharing. The specific process of the profit-sharing mechanism is shown in (4).

5.1. Contribution ratio

We define that contribution refers to the change in operation cost of each IDC when IDC i participates in coordination or not. The details can be expressed as follows:

$$\delta_i = \left| R_i^{co} - R_i^{nco} \right| \tag{4a}$$

$$\eta_i = \delta_i / \sum_{i=1}^{N^I} \delta_i \tag{4b}$$

where δ_i in (4a) represents the profits of IDC i that comes from arbitrage based on varying electricity prices across regions. R_i^{co} and R_i^{nco} represents the operation cost of IDC i when the IDC i participates in the coordination or not. η_i in (4b) represents the contribution ratio of IDC i.

5.2. Profit-Sharing problems in the coordination

The main problem of profit-sharing has two aspects. One is the profit computation in arbitrage based on varying electricity prices by the spatial–temporal flexibility of IDCs. The other is the distribution of profits in the coordination composed of geo-distributed IDCs. The detailed impact is shown as follows:

1) No coordination

The profits of the IDC not participating in the coordination can be expressed as the negative value of the power procurement cost. Therefore, the profits in the electricity market can be defined as follows:

$$R_i^{nco} = -\sum_{t \in T} (P_{it}^{nco} \gamma_{it}^{nco}) \tag{4c}$$

where P_{it}^{nco} represents the power demands of IDC i and γ_{it}^{nco} represents the locational marginal price not participating in the coordination.

2) With coordination

When the geo-distributed IDCs optimize their operation strategies in the coordinated optimization model mentioned above, they need to share the workload information in the 'third-party' company for privacy requirement. Therefore, the CSP should get some profits. For the surplus profits, they should be distributed to improve the enthusiasm and recognition of the IDCs participating the coordination. Therefore, the profits from the coordination can be defined as follows:

$$R^{CSP} = (1 - \varpi) R^{sur} \tag{4d}$$

$$R^{sur} = \sum_{i \in N^I} R_i^{co} - \sum_{i \in N^I} R_i^{nco} \tag{4e}$$

R^{sur} in (4d) and (4e) represents the cooperative surplus between the CSP and the IDCs, in which the cooperative surplus means the difference between the collective interests of the parties and the sum of the individual interests of the parties in the case of non-coordination or coordination. $1-\varpi$ represents the aggregator coefficient of the CSP, which means the proportion of the profits of the CSP in the cooperative surplus.

Workloads are migrated in spatial–temporal dimensions to make profits from the difference of LMPs by spatial–temporal flexibility of the IDCs. Therefore, the cost in the coordination can be defined as follows:

$$C^{IDC} = \sum_{t \in T} \sum_{i \in N^I} \left(P_{it}^{IDC} \gamma_{it} + \frac{1}{2} \theta^{mig} \left| P_{it}^{wor} - P_{it}^{nwor} \right| \right) - \varpi R^{sur} \tag{4f}$$

5.3. Profit-Sharing mechanism

A profit-sharing mechanism based on the Nash bargaining theory is formulated in this subsection. Many studies are inclined to apply Nash bargaining theory to share profits in a cooperation game. The existing profit-sharing mechanism concentrates on distributing the profits equally or simply according to the established agreement, in which the profits are allocated. However, each member in cooperation makes different contributions to coordinated offerings, and a rational profit-sharing mechanism should identify the different contributions of coordinated offering members and reward good behavior in coordination [34].

Therefore, this paper proposes a profit-sharing mechanism based on the asymmetric Nash bargaining method, where different contribution ratios are considered. The profit-sharing mechanism can be realized by optimizing the following Nash bargaining model:

$$\max_{\xi_i^{IDC}} \prod_{i \in N^I} \left(\xi_i - R_i^{nco}\right)^{\eta_i} \tag{4g}$$

Subject to:

$$\xi_i - R_i^{nco} \geqslant 0 \cdots\cdots \forall i \in N^I \tag{4h}$$

$$\sum_{i \in N^I} \left(\xi_i - R_i^{nco}\right) = R^{sur} \tag{4i}$$

$$R_i^{nco} = -\sum_{t \in T} P_{it}^{nco} \gamma_{it}^{nco} \tag{4j}$$

The objective function (4 g) is the Cobb-Douglas utility function, which shows that the cooperative surplus of the IDCs is apportioned to all IDCs in the coordination according to a certain weight. Constraint (4i) states that the model takes the operation cost of each IDC directly trading with the CSP as the breakdown point of Nash bargaining.

The profits are allocated to each IDC in the coordination in accordance with their contributions to the optimal operation strategy. Finally, the proposed profit-sharing mechanism can realize a distribution of profits and incentive the coordinated enthusiasm among geo-distributed IDCs.

For the above model, the analytical solution can be obtained by precise mathematical derivation. The detailed mathematical derivation can be defined as follows:

a) We find that the objective function (4 g) is nonlinear and nonconvex, and the logarithm of the objective function can be calculated. Since the logarithm function is convex and the sum of convex functions is still convex according to the properties of convex functions, the KKT condition of the objective function can be used to ensure that the optimal solution can be obtained.

$$\max_{\xi_i^{IDC}} \sum_{i \in N^I} \eta_i \ln\left(\xi_i - R_i^{nco}\right) \tag{4k}$$

b) The Lagrange function L can be defined as follows:

$$L(R) = -\sum_{i \in N^I} \eta_i \ln\left(\xi_i - R_i^{nco}\right) + \sum_{i \in N^I} \varsigma_i \left(R_i^{nco} - \xi_i\right) + \sigma_i \left(\sum_{i \in N^I} \left(\xi_i - R_i^{nco}\right) - R^{sur}\right) \tag{4m}$$

ς_i and σ_i in (4 m) represent the Lagrange multipliers of constraint (4h) and constraint (4i). The KKT conditions of the objective function (4 k) can be obtained by taking the gradient of the Lagrange function as follows:

$$\begin{cases} \dfrac{\eta_i}{\xi_i - R_i^{nco}} + \varsigma_i + \phi \sigma_i = 0 \cdots i \in N^I \\ \sum_{i \in N^I} \left(\xi_i - R_i^{nco}\right) = R^{sur} \\ \varsigma_i \left(R_i^{nco} - \xi_i\right) = 0 \\ \varsigma_i \geqslant 0 \end{cases} \tag{4n}$$

where $R_i^{nco} \neq \xi_i$, ς_i=0 can be obtained from complementary relaxation conditions (4n), and it can be substituted into Equ (4o). Equ. (4o) can be obtained as follows:

$$\frac{\eta_i}{\xi_i - R_i^{nco}} + \sigma_i = 0 \cdots i \in N^I \tag{4o}$$

From the above derivation, we can finally obtain the following result:

$$\xi_i = R_i^{nco} + \eta_i R^{sur} \tag{4p}$$

The profit-sharing results in (4p) are composed of two parts. The first part is the expected profits before the IDC participates in the coordination, and the second part considers the contribution ratio of the IDC *i* to the coordination, reflecting the flexibility value of different IDCs. In this distributed solution method, each IDC in the coordination shares the coordinated surplus fairly and reasonably.

6. Case study

It is assumed that there are three geographically distributed IDCs that coordinate workload processing and share the same basic parameters, as shown in TABLE 1 [15,33]. The initial delay-tolerant workloads of each IDC are assumed to be twice as large as the delay-sensitive workloads allocated from the front-ends [21]. To characterize the flexibility of IDCs in the spatial dimension, IDCs are set at bus 4, 16, and 22 in the IEEE 30-bus network, and the operation of the three methods is demonstrated, as depicted in Fig. 5. Moreover, wind power output at node 22 is derived from wind power rich areas in northern China, which are currently favored by IDC sites and the data of IDCs are derived from the 24-hour workload processing of independent data centers in three different regions. To demonstrate the effect of IDC's spatial–temporal flexibility potential on operation, this paper designs five scenarios for

Table 1
Parameter Setting.

Parameter	Value	Unit
$\alpha_1, \alpha_2, \alpha_3$	0.0068/0.0008/0.458	/
u	200	(CPU/h)
v	0.05	(s)
e_i^{Idle}, e_i^{Peak}	0.36/0.72	($\times 10^{-3}$MW)
T^{min}, T^{max}	30/14	(°C)
ρ	1.19	(kg/m^3)
f_i	0.2454	(m^3/h)
ε	76.53	($J{*}kg^{-1}{*}^{\circ}C^{-1}$)
P^{Cmin}, P^{Cmax}	2/20	MW

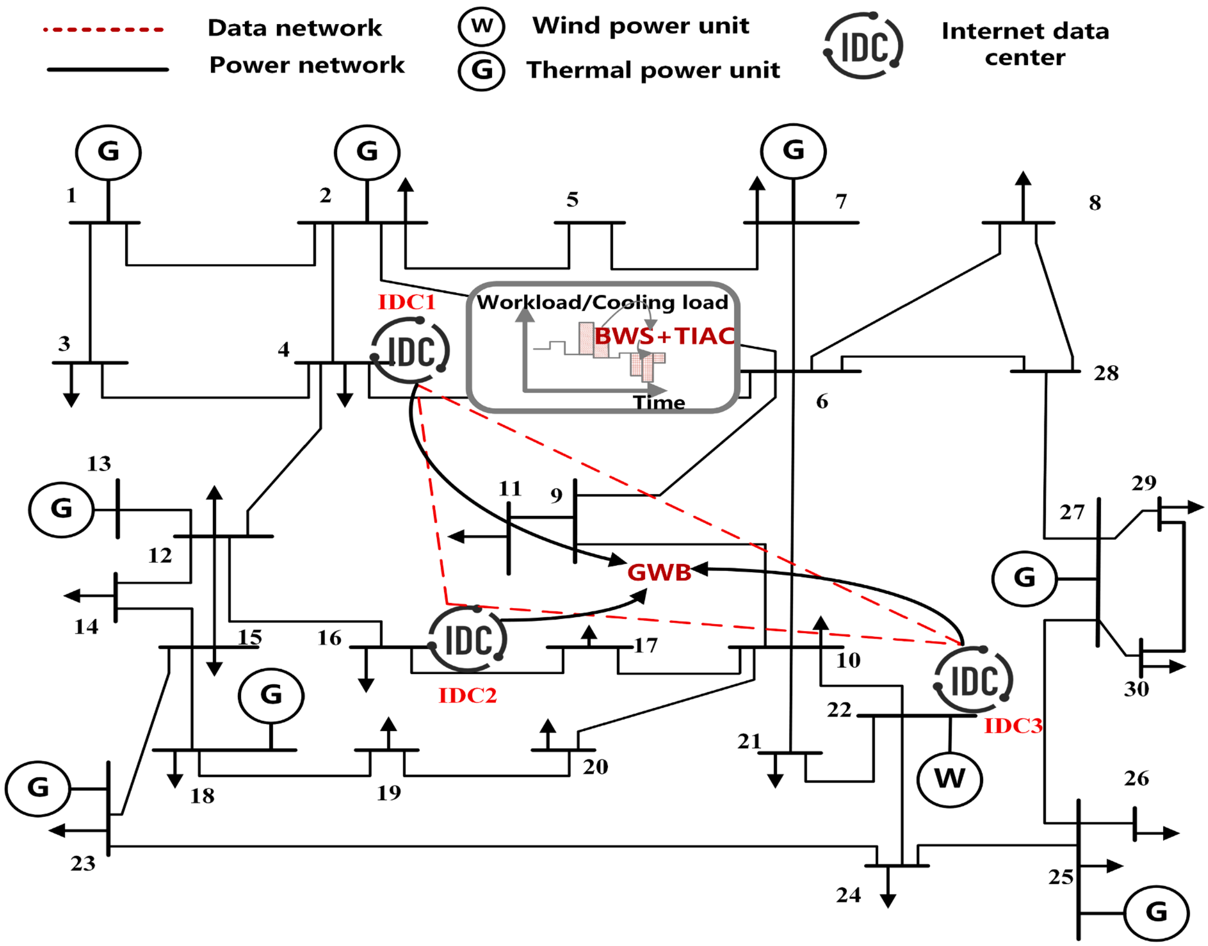

Fig 5. IEEE 30-Bus Network.

comparison in TABLE 2.

1) S1 is the proposed coordinated operation considering the energy-aware IDC loading model, which includes the workload migration and cooling system spatial–temporal flexibility provision ability of IDCs.
2) S2 refers to a contrast scenario, where only the workload migration in spatial–temporal dimension is permitted, indicating that TIAC is not considered.
3) S3 also refers to a contrast scenario, where only the flexibility potential in temporal dimension is permitted, indicating that GWB are not considered.
4) S4 also refers to a contrast scenario, where the workload migration in spatial dimension and cooling system management in temporal dimension, indicating that BWS are not considered.
5) S5 is a benchmark in which no coordination among IDCs is adopted, indicating that the spatial–temporal flexibility is not activated.

Aiming at the nonlinear term in the above coordinated optimization model, the binary variable linearization method is used to linearize the equation of absolute value (2c), the proposed model is converted into an integer programming model for solving the problem with the YALMIP toolbox [35]with 24 h as the operation timescale simulation, 1 h as an interval. The results are provided as follows.

Table 2
Five scenarios with scheduling methods.

	S1	S2	S3	S4	S5
GWB	√	√	–	√	–
BWS	√	√	√	–	–
TIAC	√	–	√	√	–

6.1. Spatial-temporal flexibility potential effect on IDC cost

To evaluate the profits of spatial–temporal flexibility potentials activated by the coordinated optimization model, The optimization results (Peak workload time slots:10o'clock-20o'clock) are shown in Fig. 6

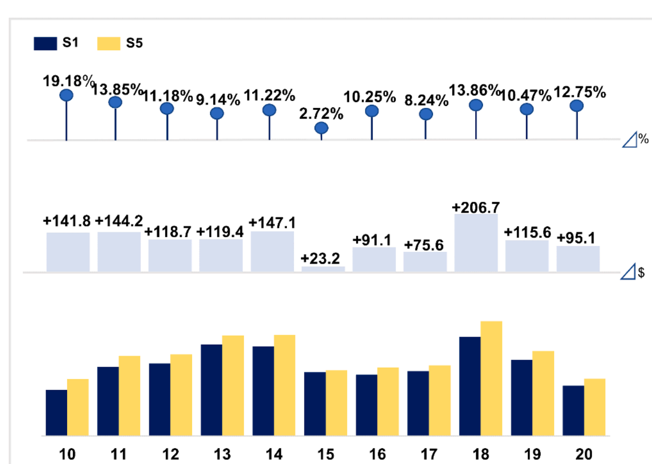

Fig. 6. The total cost of three IDCs in 10–20 time slots under the S1 and S5 scenarios.

which is associated with LMPs delivered by the lower-level model and optimal operation state showed by the upper-level model. The third layer in the figure is the cost of three IDCs, and the first layer and second layer are the percentage and the absolute value of the cost difference under the S1 and S5 scenarios.

Driven by the spatial–temporal flexibility potential, the delay-tolerant or delay-sensitive workloads can be migrated among geo-distributed IDCs or in different time slots to avoid the peak electricity price and leverage ambient temperature effect for cost reduction of IDCs. As shown in the Fig. 6, the total costs of three IDCs in S1 scenario are smaller than that in S5 scenario in 10–20 time slots to varying degrees. The following points deserves our attention. First, the cost fluctuates in waves which gradually increases in 10–14 time slots, then decreases in 15–17 time slots, peaks in the 18-time slot, and then gradually decreases again. Second, the cost decreases greatly in the time slots such as 10–12 time slots and 18–20 time slots. The reason for these points consists of two parts: the diversity of electricity prices and the energy-saving of cooling system.

The diversity of the LMP: the LMP of the node located at an IDC, such as the node 4, 16 and 22, is shown in Fig. 7. As shown in the Fig. 5, the node 22 is equipped a wind power unit. Since the operation cost of the wind power unit is much lower than that of the thermal power unit, the LMP of node 22 is also lower than other nodes with an IDC according to the definition of LMP, which leads to a large number of workloads satisfying the requirement of QoS being migrated from other two nodes to node 22 for pursuing the optimal cost such as that in the 18-time slot. That fully reflects the unique spatial flexibility of the IDCs. The energy-saving of cooling system: the decrease of power consumption of cooling system affected by the ambient temperature is shown in Fig. 8. As formulated in (1o), the ambient temperature whether indoor temperature and outdoor temperature influences cooling system management in TIAC. The effect of the flexibility simulated by TIAC on the power consumption of cooling system is positively correlated with ambient temperature, for example, the power saving caused by TIAC is highest when outdoor temperature is high in 12–14 time slots, and it decreases in 18–20 time slots, so does the outdoor temperature. However, the profits caused by LMP at the node 22 in 18–20 time slots compensate for this relatively negative effect.

Finally, according to the above coordinated optimization model, the coordinated operation cost of the IDCs in the upper-level model is reduced from 17,754$ (S5) to 15,671$ (S1) under the effect of spatial–temporal flexibility of the IDC, and its profits are 2,083$. It is verified that IDCs in the coordination equipped with multiple energy-aware scheduling methods in activating spatial–temporal flexibility potential can bring considerable profits on cost reduction.

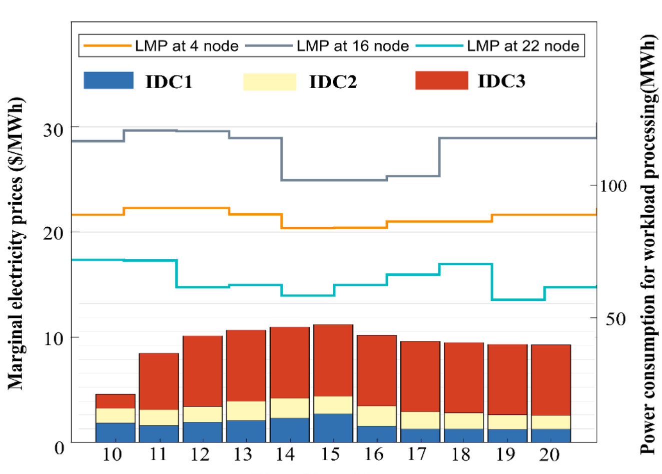

Fig. 7. The LMPs at the node 4, 16, 22 and the power consumption for workload processing of IDCs at the node 4, 16, 22.

6.2. Coupling of three scheduling methods in IDCs' coordination profits

Here, to demonstrate the advantage of the energy-aware IDC loading model, this section compares our proposed model with the conventional model that uses PUE [3] to represent power consumption, excluding IT equipment, and does not consider TIAC, such as the S2 scenario. The results are shown in the TABLE 3.

From the TABLE 3, the operation strategy given by this paper can enhance flexibility and reduce costs compared to existing research that relies solely on PUE to measure power consumption. Additionally, there may be a coupled relationship among these scheduling methods, as shown in Equ. (1o). To demonstrate the coupling of GWB, BWS, and TIAC in the flexibility potentials of geo-distributed IDCs on cost reduction, this section also designs the S2, S3 and S4 mentioned above.

Keeping the other parameters unchanged, the different scheduling methods in the four scenarios (S1, S2, S3, S4) are adopted in the above coordinated optimization model. The specific results are shown in Fig. 9. It is observed that the IDC coordinated operation costs are the lowest, and the cost curtailment is the highest when considering the three scheduling methods in coordination with spatial–temporal dimensions. This directly demonstrates the superiority of the above IDC loading model. It is noteworthy that the cost of operation can be reduced significantly by considering the temporal flexibility potentials of the cooling load scheduling method. This aspect has been overlooked in previous studies. Furthermore, if the three methods in IDCs are not coupled for cost reduction, the total cost reduction of all three IDCs over 24 time slots would be (1,014$+1,712$+1,314$)/2, which equals 2,020 $, when considering energy-aware scheduling for all GWB, BWS, and TIAC. However, the data presented in Column 1 of Fig. 9 shows that the actual cost reduction achieved by three IDCs' spatiotemporal flexibility is 2,083$, which contradicts the previous inference. This suggests that the coupling of three scheduling methods affects the IDCs' coordination profits in terms of cost reduction.

Here, if three energy-aware scheduling methods are deployed jointly in the IDC operation, the cost reduction increases by 5.05 %. This is in comparison to the cumulative cost reduction achieved when these methods are deployed separately. That is, when TIAC is introduced to GWB and BWS, in addition to the increased profits for cost reduction brought by deploying TIAC itself, significant extra improvement in cost reduction (i.e., 5.05 %) is achieved. This indicates the necessity of coordinating GWB, BWS, and TIAC in energy-aware scheduling for the optimal operation state.

6.3. Profit-Sharing results

Here, this section discusses the profit-sharing mechanism to avoid the phenomenon of incentive incompatibility and incentive the enthusiasm for an IDC's participation in the coordination. On the basis of that, the mechanism proposed in this paper can distinguish the contribution of each IDC in coordination with different contribution ratios.

Using the above simulation in Case Study A, specific results in TABLE 4 is calculated for the costs of each IDC in the S1 and S5 scenarios, as well as their contribution ratios. From the TABLE 4, it is found that after joining the coordination to schedule its workload processing decisions based on varying electricity prices across regions, IDC 3 experiences a negative cost reduction. This means that the cost optimization of IDC 3 is contradictory to that of the coordination. The reason for this incentive incompatibility is that a large amount of workloads of IDC 1 and IDC 2 are migrated to IDC 3 because of the LMPs as shown in Fig. 7, leading to a higher power consumption cost than the profits of IDC 3 brought by spatial–temporal flexibility. This will greatly discourage ID C3 from participating in coordination, so an artificial distribution of the coordination profits is necessary. If the traditional symmetric Nash bargaining method is used to distribute the coordination profits, the profits brought by flexibility potentials are evenly distributed without considering the contribution of each IDC. It is assumed that the CSP is all in

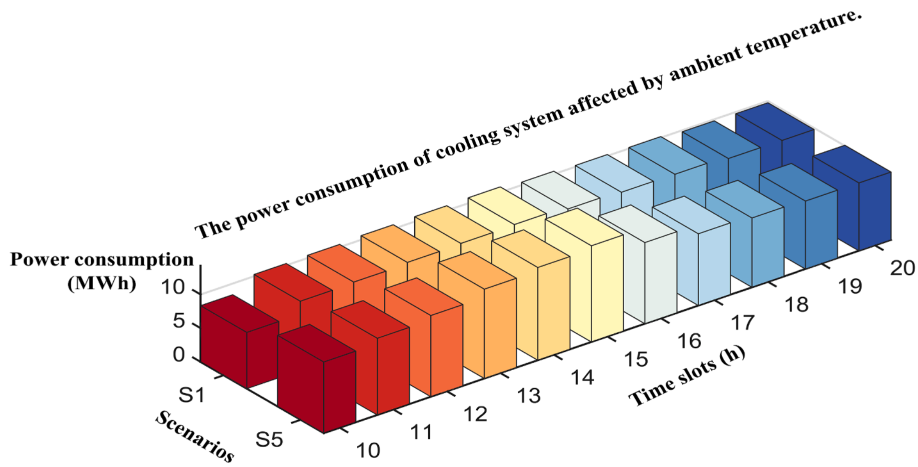

Fig. 8. The power consumption of cooling system affected by ambient temperature.

Table 3
Comparison Between S1 AND S2.

	Value ($)		Value ($)
S1	15,671	S2	16,740
S5	17,754	S5	17,754
Cost Reduction	2083	Cost Reduction	1014
Percentage	11.73 %	Percentage	5.71 %

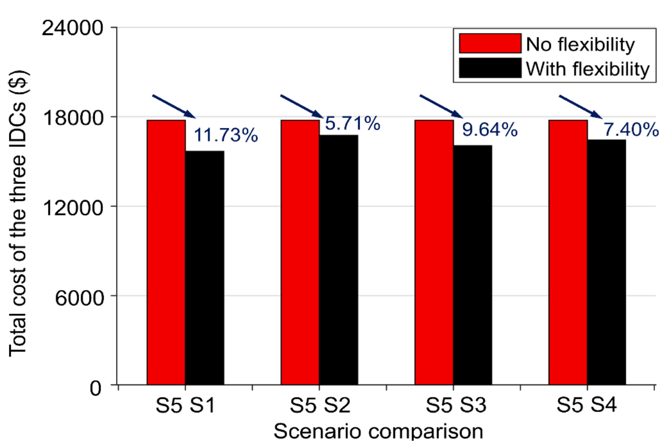

Fig. 9. Cost reduction under 4 scenario comparisons.

Table 4
Operation Costs And Contribution Ratios.

	IDC1	IDC2	IDC3
S5 ($)	6226.5	7863.1	3664.2
S1 ($)	4389.9	5959.9	5322.2
Cost Reduction ($)	1836.6	1903.2	−1658
Contribution Ratio	0.34	0.35	0.31

favor of the IDCs, that is, the profits of each data center are (2,083$/3), i.e., 694.3$, when using the symmetric Nash bargaining method to share the profits. In this extreme case, the traditional symmetric Nash bargaining method cannot compensate for the extra cost of IDC 3, which is unfair to IDC 3.

Therefore, the coordination profits are distributed using the asymmetric Nash bargaining method based on the contribution ratios in this paper. Firstly, the CSP is allocated 5 % of the coordination profits (ϖ=0.95, in (4d)) for its aggregator status. Secondly, the extra cost of IDC 3 is compensated for, at least to restore its costs to the level of the S5 scenario. The profits from coordination are distributed based on the contribution ratios. It is important to note that the profit-sharing mechanism proposed in the paper is not applicable in scenario S3, where there is no coordination among geo-distributed IDCs. The final allocation results for S1, S2 and S4 scenario are presented in Fig. 10.

It can be seen that the cost of each IDC decreased due to the asymmetric Nash bargaining method. Among them, the cost of IDC3 has the largest decrease compared with the other two IDCs because the locational marginal price of the node where IDC3 is located is lower than that of the other two IDCs. Therefore, when the coordinated model optimizes the operation strategy of the IDCs, a large number of workloads will be migrated to IDC3 for processing. It is a good counterpart to the law 'You do the most, you get the most', and it is verified that the profit-sharing mechanism is reasonable and fair. In addition, in S4 scenario, the extra costs of IDC 3 are more than the profits caused by the flexibility potential, which further affirms the superiority of the proposed optimization method in S1 scenario.

7. Conclusions and discussions

The deregulation of the electricity market can facilitate coordination optimization between geo-distributed internet data centers through market instruments, resulting in a more economically viable power consumption strategy. This paper proposes an energy-aware coordinated operation strategy of geographically internet data centers. The proposed approach has several advantages over alternative methods.

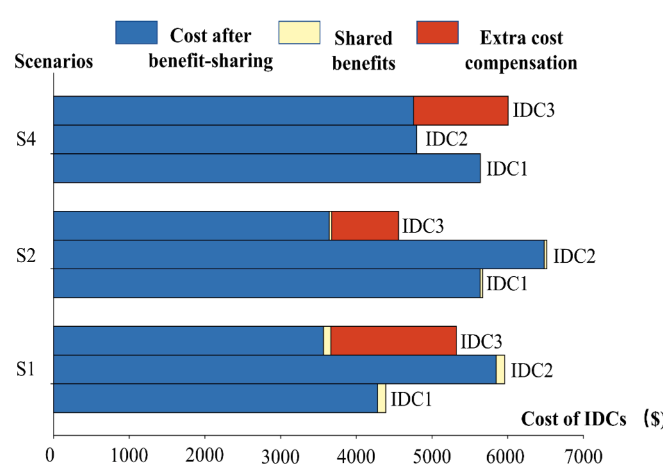

Fig. 10. The final profit-sharing results.

Firstly, it introduces an energy-aware internet data center loading model that considers coupled scheduling methods, which enhances the flexibility of geo-distributed internet data centers. Secondly, the electricity market decision-making model has transitioned from a conventional approach based on price followers to a bi-level model that considers the game interactions of multiple subjects to achieve profits through arbitrage from multiple electricity prices. Thirdly, a reasonable and fair profit-sharing mechanism has been designed to incentivize participation in coordination and ensure the interests of each internet data center are protected. In the case studies, five cases are compared, which demonstrates following.

1) The issue of high operational costs in internet data centers has garnered significant attention from the academic community. Without economic incentives, the cost of power procurement and the added operational costs related to flexibility have dissuaded many internet data center operators, leading to a waste of resources in terms of the potential for spatiotemporal internet data centers. The coordination operation of geo-distributed internet data centers proposed in this paper can be performed successfully to lower their coordinated operation costs with varying electricity prices. In our case, the total costs of the operation decreased by 11.73 %.
2) Three energy-aware scheduling methods should be considered jointly rather than separately, and the operations of these scheduling methods should be implemented simultaneously to exploit more flexibility potentials in coordination among multiple internet data centers. Through a case study, the coupling among the three methods is empirically analyzed, and it is verified that the coupling relationship will positively impact coordinated profits on the cost reduction. Simulation results show that when these three methods are deployed together, by comparison with the summation of the coordination profits when the methods are deployed separately, the profits increase by 5.05 % from the economic perspective.
3) To fully utilize local resources, geo-distributed internet data centers tend to assemble a coordination. However, in the process of coordination, individual interests may be contrary to collective interests, so a reasonable profit-sharing mechanism is necessary. Then, the concept of contribution ratios is introduced to express the flexibility value of each internet data center. Based on the contribution ratios, the proposed profit-sharing mechanism distributes these coordination profits to each internet data center for cost reduction. In the case study, the internet data centers3 gains the most coordination profits because the most workload it processes, which affirms the rationality of profit-sharing mechanism.

8. Limitations and future work

In the context of the electricity market, it is of great theoretical and practical significance to study the optimal operational strategy of geo-distributed IDCs. This paper specifically examines the participation of 'IDCs' in electricity market transactions, with a focus on energy and information interactions, as they carry service requests from cloud users. However, it is important to note that the power grid serves not only as a carrier for various information and communication technologies, but also as a significant information flow generation and convergence center. If considering the data center serves as a service provider for grid information, a deeper coordination between information flow and power flow should be achieved. Therefore, further in-depth research is necessary on the coordination of data centers within the power system, specifically in terms of grid planning, operation, and control.

CRediT authorship contribution statement

Shibo Zhou: Formal analysis, Investigation, Methodology. **Ming Zhou:** Funding acquisition, Supervision, Writing – review & editing. **Zhaoyuan Wu:** Methodology. **Yuyang Wang:** Software. **Gengyin Li:** Data curation, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported in part by the Joint Funds of the National Natural Science Foundation of China (No. U22B20110), and in part by the National Natural Science Foundation of China (No. 52207098).

Appendix

A. Derivation of locational marginal price.

First, we can abstract the equations (3) in the lower level into the following mathematical model:

$$\min f(x) \tag{A1}$$

$$g(x) = b \tag{A2}$$

The objective function (A1) is the abstract form of equation (3b), which represents the operation cost. The constraint (A2) is the abstract form of the whole system operation constraints. x and b is the abstract form of the electric outputs and loads, respectively.

Therefore, when solving the optimization model (A), the Lagrange multiplier is the locational marginal price. The specific derivation process is as follows:

(a) The Lagrange function V can be defined as follows:

$$V = \min f(x) + \gamma(b - g(x)) \tag{A3}$$

In constraint (A3), we assume that b is the variable, then we take (x^*, λ^*) as the minimum value. So, the optimal solution can be expressed as ($x^*(b)$, $\lambda^*(b)$).

(b) The Lagrange function $V(b)$ can be defined as follows:

$$V(b) = f(x^*(b)) + \lambda^*(b)[b - g(x^*(b))] \tag{A4}$$

(c) The derivation of variable b with respect to Eq. (5d) can be defined as follows:

$$\begin{aligned}V'(b) &= f_x(x^*(b))\frac{\partial x^*(b)}{\partial b} + \frac{\partial \lambda^*(b)}{\partial b}[b - g^*(x(b))] + \lambda^*(b)[1 - g_x((x^*(b))\frac{\partial x^*(b)}{\partial x(b)}]\\ &= [f_x(x^*(b)) - \lambda^*(b)g_x((x^*(b))]\frac{\partial x^*(b)}{\partial x(b)} + \frac{\partial \lambda^*(b)}{\partial b}[b - g^*(x(b))] + \lambda^*(b)\end{aligned} \tag{A5}$$

(d) Based on the optimal solution ($x^*(b)$, $\lambda^*(b)$), we can conclude as follows:

$$f_x(x^*(b)) - \lambda^*(b)g_x((x^*(b)) = 0 \tag{A6}$$

$$b - g^*(x(b)) = 0 \tag{A7}$$

(e) Then, we combine Eq. (5f) and Eq. (5 g) to get the result.

$$V'(b) = \lambda^*(b) \tag{A8}$$

Therefore, the Lagrange multiplier λ is the marginal value of the variable b, for that, how much the optimal value of the function $f(x)$ increases when b increases by one unit. This can prove that the locational marginal price $\gamma_{i,t}$ in (2b) is the dual variable of constraint (3f), which we also call it 'shadow price'.

B. Derivation of linearization

Due to the absolute value in the upper-level model, we need to make the linearization with the help of 0–1 binary variables. The specific derivation is as follows:

(a) The parameters can be defined as follows:

$$\begin{cases} G_i = \sum_{t\in T} P_{it}^{wor} \\ G_i^n = \sum_{t\in T} P_{it}^{nwor} \\ f_i = \left|\sum_{t\in T} P_{it}^{wor} - \sum_{t\in T} P_{it}^{nwor}\right| = \left|G_i - G_i^n\right| \end{cases} \tag{B1}$$

(b) The piecewise function can be defined as follows:

$$f_i = \begin{cases} G_i - G_i^n G_i^n \leqslant G_i \leqslant M \\ G_i^n - G_i 0 \leqslant G_i \leqslant G_i^n \end{cases} \tag{B2}$$

In constraint (B2), M is an infinite number. The piecewise points of this function f_i are as follows:

$$\begin{cases} f(b_1) = G_i^n b_1 = 0 \\ f(b_2) = 0 b_2 = G_i^n \\ f(b_3) = M b_3 = M \end{cases} \tag{B3}$$

(c) When we introduce the new variables w_{ik} and z_{ik}, the function f_i can be defined as follows:

$$G_i = \sum_{k=1}^{3} w_i b_i \tag{B4}$$

$$f_i = \sum_{k=1}^{3} w_i f(b_i) \tag{B5}$$

w_{ik} is the continuous variable and z_{ik} is the 0–1 binary variable. They need to meet the following constraints to complete linearization:

$$\begin{cases} w_1 \leqslant z_1 \\ w_2 \leqslant z_1 + z_2 \\ w_3 \leqslant z_2 \\ w_1 + w_2 + w_3 = 1 \\ z_1 + z_2 = 1 \end{cases} \tag{B6}$$

Adding constraints (B4) – (B6), this model can be transformed into a mixed integer linear programming model.

References

[1] Wu Z, Wang J, Zhong H, Gao F, Pu T, Tan C-W, et al. Sharing Economy in Local Energy Markets. J Mod Power Syst Clean Energy 2023;11(3):714–26.

[2] Cao Y, Cao F, Wang Y, Wang J, Wu L, Ding Z. Managing data center cluster as non-wire alternative: A case in balancing market. Appl Energy 2024;360:122769.

[3] Wu Z, Chen L, Wang J, Zhou M, Li G, Xia Q. Incentivizing the Spatiotemporal Flexibility of Data Centers Toward Power System Coordination. IEEE Trans Network Sci Eng 2023;10(3):1766–78.

[4] Cao Y, Chen M, Zang S. Data-driven flexibility assessment for internet data center towards periodic batch workloads. Appl Energy 2022;324:119665.

[5] Niu T, Hu B, Xie K, Pan C, Jin H, Li C. Spacial coordination between data centers and power system considering uncertainties of both source and load sides. Int J Electr Power Energy Syst 2021;124:106358.

[6] Ni J, Bai X. A review of air conditioning energy performance in data centers. Renewable Sustainable Energy Rev 2017;67:625–40.

[7] V. S. Simon, A. Siddarth, D. Agonafer. Artificial Neural Network Based Prediction of Control Strategies for Multiple Air-Cooling Units in a Raised-floor Data Center. 2020 19th IEEE Intersociety Conference on Thermal and Thermome-chanical Phenomena in Electronic Systems (ITherm), Orlando, FL, USA, 2020, 334-340.

[8] Yuan H, Bi J, Zhang J, Zhou M. Energy Consumption and Performance Optimized Task Scheduling in Distributed Data Centers. IEEE Trans Syst Man Cybern: Syst 2022;52(9):5506–17.

[9] Chen B, Che Y, Zheng Z, Zhao S. Multi-objective robust optimal bidding strategy for a data center operator based on bi-level optimization. Energy 2023;269:126761.

[10] Chen M, Gao C, Shahidehpour M, Li Z. Incentive-Compatible Demand Response for Spatially Coupled Internet Data Centers in Electricity Markets. IEEE Trans Smart Grid 2021;12(4):3056–69.

[11] Aydin H, Melhem R, Mosse D, Mejia-Alvarez P. Power-aware scheduling for periodic real-time tasks. IEEE Trans Computers 2004;53(5):584–600.

[12] Hajiamini S, Shirazi B, Crandall A, Ghasemzadeh H. A Dynamic Programming Framework for DVFS-Based Energy-Efficiency in Multicore Systems. IEEE Trans Sustainable Comput 2020;5(1):1–12.

[13] Wang H, Wang Q, Tang Y, Ye Y. Spatial load migration in a power system: Concept, potential and prospects. Int J Electr Power Energy Syst 2022;140:107926.

[14] Hossain MS, Rahaman S, Kor A-L, Andersson K, Pattinson C. A Belief Rule Based Expert System for Datacenter PUE Prediction under Uncertainty. IEEE Trans Sustainable Comput 2017;2(2):140–53.

[15] Chen M, Gao C, Shahidehpour M, Li Z, Chen S, Li D. Internet Data Center Load Modeling for Demand Response Considering the Coupling of Multiple Regulation Methods. IEEE Trans Smart Grid 2021;12(3):2060–76.

[16] Hu X, Li P, Sun Y. Minimizing Energy Cost for Green Data Center by Exploring Heterogeneous Energy Resource. J Mod Power Syst Clean Energy 2021;9(1):148–59.

[17] Li W, Qian T, Zhao W, Huang W, Zhang Y, Xie X, et al. Decentralized optimization for integrated electricity–heat systems with data center based energy hub considering communication packet loss. Appl Energy 2023;350:121586.

[18] Lyu J, Zhang S, Cheng H, Yuan K, Song Y, Fang S. Optimal Sizing of Energy Station in the Multienergy System Integrated With Data Center. IEEE Trans Ind Appl 2021;57(2):1222–34.

[19] Yu L, Jiang T, Zou Y. Price-Sensitivity Aware Load Balancing for Geographically Distributed Internet Data Centers in Smart Grid Environment. IEEE Trans Cloud Comput 2018;6(4):1125–35.

[20] Jiang C, Tseng C, Wang Y, Lan Z, Wen F, Chen F, et al. Optimal Pricing Strategy for Data Center Considering Demand Response and Renewable Energy Source Accommodation. J Mod Power Syst Clean Energy 2023;11(1):345–54.

[21] Wang P, Cao Y, Ding Z, Tang H, Wang X, Cheng M. Stochastic Programming for Cost Optimization in Geographically Distributed Internet Data Centers. CSEE J Power Energy Syst 2022;8(4):1215–32.

[22] Jian J, Zhao J, Ji H, Bai L, Xu J, Li P, et al. Supply Restoration of Data Centers in Flexible Distribution Networks with Spatial-Temporal Regulation. IEEE Trans Smart Grid 2023;3286844.

[23] Zhang G, Zhang S, Zhang W, Shen Z, Wang L. Distributed Energy Management for Multiple Data Centers With Renewable Resources and Energy Storages. IEEE Trans Cloud Comput 2022;10(4):2469–80.

[24] Kumar N, Aujla GS, Garg S, Kaur K, Ranjan R, Garg SK. Renewable Energy-Based Multi-Indexed Job Classification and Container Management Scheme for Sustainability of Cloud Data Centers. IEEE Trans Ind Inf 2018;15(5):2947–57.

[25] Chen M, Gao C, Shahidehpour M, Li Z. Proliferation of Small Data Networks for Aggregated Demand Response in Electricity Markets. IEEE Trans Power Syst 2022;37(3):2297–311.

[26] Zhou Z, Liu F, Li Z. Bilateral Electricity Trade Between Smart Grids and Green Datacenters: Pricing Models and Performance Evaluation. IEEE J Sel Areas Commun 2016;34(12):3993–4007.

[27] Lu X, Zhang P, Li K, Wang F, Li Z, Zhen Z, et al. Data Center Aggregators' Optimal Bidding and Benefit Allocation Strategy Considering the Spatiotemporal Transfer Characteristics. IEEE Trans Ind Appl 2021;57(5):4486–99.

[28] Wahlroos M, Pärssinen M, Rinne S, Syri S, Manner J. Future views on waste heat utilization – Case of data centers in Northern Europe. Renewable Sustainable Energy Rev 2018;82:1749–64.

[29] Jin C, Bai X, Yang C, Mao W, Xu X. A review of power consumption models of servers in data centers. Appl Energy 2020;265:114806–23.

[30] Guo Y, Li H, Pan M. Colocation Data Center Demand Response Using Nash Bargaining Theory. IEEE Trans Smart Grid 2018;9(5):4017–26.

[31] Rao DMKKV, Ukil A. Modeling of Room Temperature Dynamics for Efficient Building Energy Management. IEEE Trans Syst Man Cybern: Syst 2020;50(2):717–25.

[32] Gu L, Hu J, Zeng D, Guo S, Jin H. Service Function Chain Deployment and Network Flow Scheduling in Geo-Distributed Data Centers. IEEE Trans Network Sci Eng 2020;7(4):2587–97.

[33] Wu Z, Wang J, Zhou M, Xia Q, Tan C-W, Li G. Incentivizing Frequency Provision of Power-to-Hydrogen toward Grid Resiliency Enhancement. IEEE Trans Ind Inf 2023;19(9):9370–81.

[34] Wu C, Zhou D, Lin X, Wei F, Chen C, Ma Y. A novel energy cooperation framework for community energy storage systems and prosumers. Int J Electr Power Energy Syst 2022;134:107428.

[35] J. Lofberg. YALMIP: a toolbox for modeling and optimization in MATLAB. 2004 IEEE International Conference on Robotics and Automation (IEEE Cat. No.04CH37508), Taipei, Taiwan, 2004, 284-289.