European Journal of Engineering and Technologies

Vol. 2 No. 1 2026

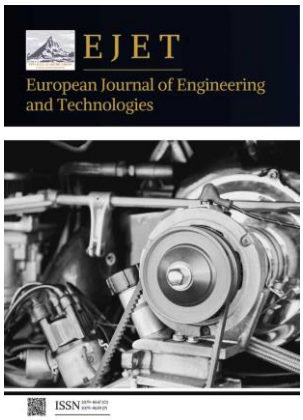
Article **Open Access**

Energy efficiency and sustainability strategies for data centers

Yuhan Zhou [1,*]

[1] Department of Economics, Virani Undergraduate School of Business, Rice University, Houston, 77005, United States

* Correspondence: Yuhan Zhou, Department of Economics, Virani Undergraduate School of Business, Rice University, Houston, 77005, United States



Abstract: With the development of informatization, the energy consumption level of data centers is increasing day by day. How to transform towards a more efficient and greener direction has thus become a technical issue for its development. This article conducts a detailed analysis of the composition of energy efficiency consumption in data centers, and discusses technology-oriented sustainable development technologies and efficient energy solutions. Finally, on this basis, the implementation methods of various energy-saving technologies such as server energy conservation, virtualization, and cooling systems, as well as the sustainable development strategy combining renewable energy, prefabricated modularization, and climate synergy, are discussed, providing theoretical and methodological support for the construction of green data centers.

Keywords: data center; energy efficiency; sustainable development; virtualization technology; clean energy

Received: 10 December 2025
Revised: 03 February 2026
Accepted: 16 February 2026
Published: 22 February 2026

1. Introduction

Influenced by the large-scale investment in cloud computing, big data and intelligent technologies, data centers have become an important infrastructure of modern information technology, and the energy consumption issue of data centers has become a major problem [1]. Its huge energy consumption not only leads to an increase in operating costs, but also brings about the adverse consequence of environmental pollution. It is urgent to propose systematic solutions to alleviate the contradiction between energy utilization and environmental protection. This article, through the discussion of technical solutions such as the energy structure adjustment of data centers, efficient energy-saving technologies and sustainable strategies, aims to build green data centers [2].

2. The theoretical basis of energy efficiency and sustainable development

2.1. The basic composition of energy consumption in data centers

IT equipment is the main part of energy consumption in data centers, including servers, storage devices, network servers, etc. These are all major energy consumers. The cooling system in data centers also accounts for a certain proportion, providing a safe operating environment for equipment. Commonly used cooling technologies include air cooling, water cooling, and indirect evaporative cooling, etc. They have characteristics such as continuous operation and high power consumption, which have a great impact on the PUE of data centers. In the power supply and conversion section, there are various devices such as UPS, transformers, PDUs, etc. During the transmission process, certain losses are bound to occur.

With the development of high-performance computers and large-scale data centers, the computing density has been increasing, and thus the energy consumption level has risen. Traditional data centers with stacked hardware networking and extensive cooling technologies can no longer achieve the basic goal of balancing business development and environmental protection [3].

2.2. Technology-driven model for Sustainable Development

Cloud computing data centers not only require high-efficiency computing power, storage capacity and transmission capacity, but also need to meet the relative balance of carbon consumption, environmental protection and resource consumption [4]. It is necessary to build a comprehensive management optimization model driven by technology and with unified consideration of multi-level indicators. It can be expressed mathematically as the following objective function:

$$\min_{x}\left(E(x)+\lambda\cdot C(x)\right) \quad (1)$$

Among them, $E(x)$ represents the total energy consumption of the data center under configuration x, $C(x)$ represents the corresponding carbon emissions or environmental costs, and λ is the trade-off coefficient between the two. By adjusting the λ values, a strategic balance can be achieved between pursuing low energy consumption and low carbon emissions [5].

2.3. Technical Principles and Modeling Methods of Efficient energy systems

The establishment of large-scale energy systems requires modeling and time-varying control of multiple physical quantities such as energy flow, heat flow, and power flow [6]. For instance, when it comes to server cooling equipment, the working efficiency of the cooler is closely related to the cooling heat load of the data center. The thermodynamic model can well represent the heat exchange behavior of the cooling medium:

$$Q=m\cdot c\cdot \Delta T \quad (2)$$

Among them, Q represents the transferred heat, m is the mass flow rate of the cooling medium, c is its specific heat capacity, and ΔT is the temperature difference between the inlet and outlet of the cooling system. Mass flow rate and temperature difference regulation are effective ways to increase cooling capacity and reduce energy consumption. When conducting system modeling, the integrated setting of multiple variables can also be adopted with the aid of cybernetics and simulation evaluation. For instance, by adopting Model Predictive Control (MPC), the cooling capacity can be adjusted in real time while maintaining a constant temperature in the data center to achieve energy-saving effects, reduce the overall consumption of the system, and ensure the energy efficiency and availability of the data center [7].

3. Core technical paths for enhancing energy efficiency in data centers

3.1. Energy-saving optimization technologies at the hardware level

The source of energy consumption in data centers is hardware equipment. Nowadays, server hosting providers generally start to recommend the use of cpus with lower power consumption, such as those using FinFET process cpus or SoCs, and ARM-based cpus. Compared with x86, these can achieve lower power consumption under the same computing power [8]. Meanwhile, as more and more data centers gradually replace mechanical hard drives with solid-state drives to reduce storage load consumption and access energy consumption, this development trend is particularly significant in read/write I/ O-intensive work applications. In terms of power input, industry experts generally recommend power equipment with efficient conversion, such as power modules of 80 PLUS Platinum or Titanium grade, which can achieve a power of more than 94% under normal load conditions, achieving efficient energy conservation and consumption reduction while reducing the load of UPS and the cooling system [9].

To more accurately explore the potential for energy conservation, the existing server systems all support the DVFS function, which can dynamically set the operating frequency and supply voltage of the CPU core according to the status of the workload [10]. If the workload is relatively light, the frequency of the CPU core can be set to the minimum value, which greatly reduces the static power consumption. In addition, the server is equipped with a dedicated hardware module for monitoring energy consumption (such as Intel RAPL), which can obtain and save energy consumption information in milliseconds, enabling the high-level management system to obtain the energy consumption information of each core and memory within each millisecond time period in real time, which is conducive to making more reasonable decisions (as shown in Figure 1).

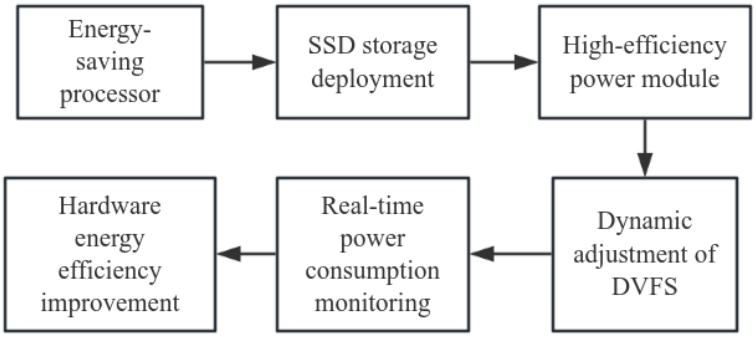


Figure 1. Flowchart of the hardware energy-saving optimization technology path.

3.2. Resource Scheduling and Virtualization Control System

The core of the energy efficiency control system in the new type of data center is to rely on virtualized resource scheduling and optimization management technology [11]. The basic idea is to implement logical abstraction and flexible and detailed management at the physical device layer, thereby improving the utilization rate of computing resources in the time dimension and the space dimension, and ultimately achieving the effect of energy conservation. Typically, data centers choose to use virtualization platforms such as KVM, VMware ESXi, or Xen to divide the CPU, RAM, and I/O resources of servers into several logical units for use by virtual machines or containers. To achieve this high-concurrency scheduling effect, a series of resource managers such as Kubernetes and OpenStack Nova Scheduler are necessary [12]. They can dynamically collect and model the resource conditions of all machines in the entire data center. For example, the CPU usage rate, memory occupancy rate, power consumption, network traffic usage, etc. of each machine.

Multi-objective optimization is generally the basic principle of energy efficiency optimization scheduling strategies. The scheduler adopts heuristic algorithms, linear programming methods or reinforcement learning to solve resource allocation. The common objective functions are as follows:

$$min\left(\sum_{i=1}^{N} P_i(u_i) + \alpha \cdot D_{mig} + \beta \cdot \Delta T\right) \quad (3)$$

Among them, $P_i(u_i)$ represents the power consumption model of the i-th node under load u_i, D_{mig} represents the system performance disturbance cost caused by virtual machine migration, ΔT is the SLA violation time window during the migration process, and α and β are the weight parameters. Both of the above two types of models can achieve dynamic scheduling decisions through methods such as particle swarm optimization (PSO), simulated annealing or deep reinforcement learning. During operation, the scheduler responds to the energy usage model, decides to schedule tasks to a small number of high-load nodes for execution, and sends sleep commands to those nodes with light loads or those that have not worked for a long time, enabling them to enter a low-power mode of C6 level or higher precision (such as S3). In addition, the

scheduling system needs to have the ability of online migration to ensure flexibility and business continuity. The migration is achieved through the incremental memory page synchronization mechanism, pre-copy technology and network I/O redirection. To reduce migration energy consumption and service interruption time, many systems have introduced bandwidth prediction mechanisms and I/O speed control mechanisms.

Furthermore, after container technology was widely accepted by the general public, the scheduling function was developed to enable resource allocation constraints and automatic scaling (HPA) at the pod level. By using the data collected by monitoring tools such as Prometheus and Grafana for automatic adjustment, on this basis, not only can the real-time load changes be guaranteed, but also the optimal energy consumption status of nodes can be met (as shown in Figure 2).

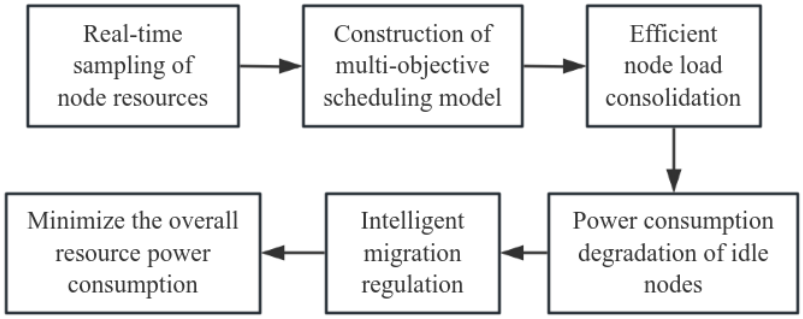


Figure 2. Energy efficiency optimization flowchart of the resource scheduling and virtualization control system.

3.3. Cooling and Environmental Control System

The thermal load control of the data center can be regarded as a typical real-time feedback control system, and its core lies in the closed loop among environmental temperature perception, modeling and end effectors. The air conditioning equipment first collects information on temperature, wet bulb temperature, air volume and cold channel pressure difference by deploying a large number of distributed sensors (such as DS18B20, DHT22, etc.) and sends it to the host BMS through the Modbus TCP/IP protocol.

After integrating the data, the thermal trend of the time series can be predicted through CNN or RNN, such as historical load, outer ring temperature, and the intensity of temperature drop in the zone, etc. This can be used to control the outlet water temperature of the chiller (6 ~ 12℃) and the PWM frequency of the fan, achieving feedback on efficiency during the regulation process. The cooling system is the contact between the chip and the cold plate for heat exchange. The heat exchange efficiency model is shown as follows:

$$Q = h \cdot A \cdot (T_s - T_f) \tag{4}$$

Here, Q is the heat transfer per unit time, h is the thermal conductivity coefficient, A is the contact area, T_s and T_f are the temperatures of the heat source surface and the cooling liquid respectively. This system will monitor and calculate the values and load parameters in real time to achieve flexible control of the flow rate or changes in heat dissipation, so as to achieve a better heat dissipation effect. When integrating the system, BACnet or OPC-UA technologies are adopted to control the air conditioning units, mufflers, chilled water, and waste heat recovery equipment, forming an intelligent and controlled building environment energy management and control platform, in order to achieve the goals such as the total energy consumption, carbon emissions, and maintenance costs of the system (as shown in Figure 3).

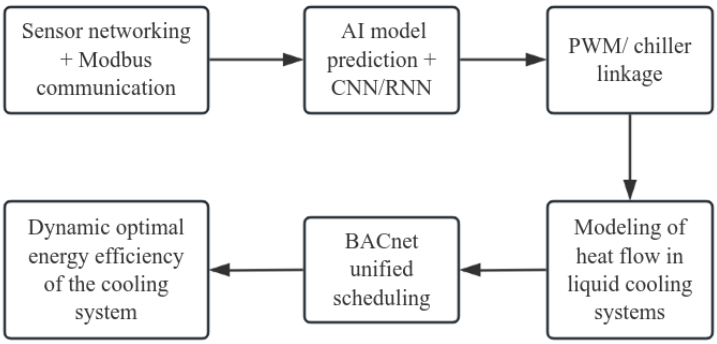


Figure 3. Flowchart of the cooling and Environmental control system.

4. Build a sustainable data center strategy

4.1. Clean Energy and Power System Integration

The goals of "carbon peaking and carbon neutrality" have promoted the application of clean energy that cannot be ignored in the sustainable development of data centers, especially in the infrastructure construction of the energy structure transformation. The most common clean energy integration methods in data centers include solar photovoltaic, wind power, geothermal and hydrogen energy systems, which can be classified into three types: On-site, Near-site PPA, and long-distance green power markets (RECs, also known as green certificates).

In order to realize the vision of making full use of various energy sources, a microgrid structure is adopted in the design process of the data center, and an energy management system is utilized to regulate the mutual flow of various energy sources, including renewable energy, batteries, and the public power grid. This energy management logic of the data center microgrid is mainly applied to model-based predictive control or rolling optimization, based on photovoltaic cells, solar radiation, wind speed and load power, and then estimates are made to formulate the optimal energy ratio.

As a standard accessory for the grid-connected application of renewable energy, the energy storage system is composed of units such as lithium iron phosphate battery packs, flow batteries, and supercapacitors. The battery management system (BMS) collects data such as voltage, current, and temperature to customize charging and discharging strategies, thereby achieving an "imbalance" between the output and use of renewable energy. In higher-level systems, energy exchange as energy storage management has been introduced into the power system, such as implementing flexible regulation in the form of electricity prices, carbon emission costs, and even energy efficiency credits. For instance, energy storage devices can provide power during peak electricity consumption periods to achieve peak shaving capacity, and sell excess renewable energy power at the lowest grid rate to save costs and protect the environment. The realization of such supporting technologies and the construction of a low-carbon, high-efficiency and adaptive energy system platform mean that data centers have stronger capabilities in energy control, environmental adaptability and economic operation, which is precisely the support of green data centers.

4.2. Edge Data Center and Modular Design

With the vigorous development of technologies such as 5G, the Internet of Things, and AI, data manufacturing is evolving from centralized and diversified to more decentralized and geographically oriented, which has led to a shift from the model based on large-scale centralized data centers to more edge data centers (EDCs). These EDCs, which feature nearby services, low response latency and flexible configuration, provide efficient support for computationally intensive services or real-time services, such as intelligent manufacturing, video analysis and autonomous vehicles. Their main purpose

is to utilize distributed data computing capabilities to alleviate the load on centralized servers and enhance the overall energy efficiency and reliability of the system.

Modular design is the core science and technology for the implementation of edge data centers, which is composed of pre-designed standardized components to form prefabricated modules, including IT (such as servers and data storage), cooling (such as air or liquid cooling), power supply (such as uninterruptible power supply UPS and distribution), and management systems (such as monitoring and networking). Modular data centers can be installed and configured in a short period of time to adapt to changing demands. Through prefabrication and on-site assembly, the establishment time can be reduced to just a few weeks, significantly lowering the construction investment and testing costs. In addition, by effectively handling the space layout within the module and the energy design of the cooling system, the operating energy consumption is reduced by increasing the heat transfer efficiency and PUE.

In terms of technology, edge data centers are typically equipped with local EMS in conjunction with renewable energy systems such as solar and wind power generation, and AI is adopted to achieve integrated energy efficiency management for equipment load sharing, heat dissipation regulation, and fault prediction. At the same time, it can also be logically interconnected through SDN and container scheduling systems (such as Kubernetes) to form a computing resource pool that can be elastically allocated, enhancing the end-to-end system reliability and disaster recovery capabilities (See Table 1).

Table 1. Comparison between Centralized Data Centers and Modular Edge Data Centers.

Project	Centralized data center	Modular edge data center
Deployment cycle	Long (6 to 24 months)	Short (2 to 6 months)
Initial investment	High	Medium to low
Scalability	Weak	Strong (expanded by module)
Energy utilization efficiency	Medium (depending on the scale)	High (fine control within the module)
Clean energy adaptability	Limited	Easy to adapt locally
Network delay	High (away from the user)	Low (close to the user)
The foundation of resource control technology	Traditional centralized scheduling	SDN + Edge control
Typical application scenarios	Cloud computing, data warehouse	IoT, industrial Internet, Internet of Vehicles

4.3. Construction of the Energy and Environment Collaborative Supervision System

To meet the requirements of intelligence and environmental protection in data centers, it is necessary to replace the original simple and separate energy consumption monitoring mode. By building an integrated management system that can be connected and jointly managed with the environment, the continuous operation of data centers can be supported, enabling them to have functions such as real-time data collection, multi-perspective data analysis, intelligent decision-making, and automated execution that meet the needs of data centers. It is also necessary to achieve centralized management with the underlying systems involved, such as energy, power, air conditioning systems, carbon dioxide emissions, and environmental quality, that is, to carry out vertical integration of data.

From a technical perspective, the collaborative monitoring system is mainly divided into three major modules. The deployment of perceptrons and IoT devices collects basic information such as energy usage, current, pressure, temperature, humidity, air quality, and refrigerant flow rate, featuring high frequency and high precision. Data analysis and modeling, based on artificial intelligence and big data technologies, have developed

models for energy usage prediction, carbon dioxide emission estimation, and reverse control of freezing effects. Commonly used techniques include time series methods (such as LSTM), multivariate linear regression, and deep reinforcement learning, etc. The final step is the actuator of the operation, such as an energy management system or a data center infrastructure management system (DCIM). This actuator can automatically schedule cooling, load migration and power capacity based on the analysis results, thereby achieving multi-system-level coordination and real-time optimal control.

Secondly, this supervision system should have an operation interface and an alarm system, which enables maintenance personnel to obtain energy consumption information of the data center in a timely manner, handle emergencies promptly, and improve the efficiency and intelligence level of energy management (as shown in Figure 4).

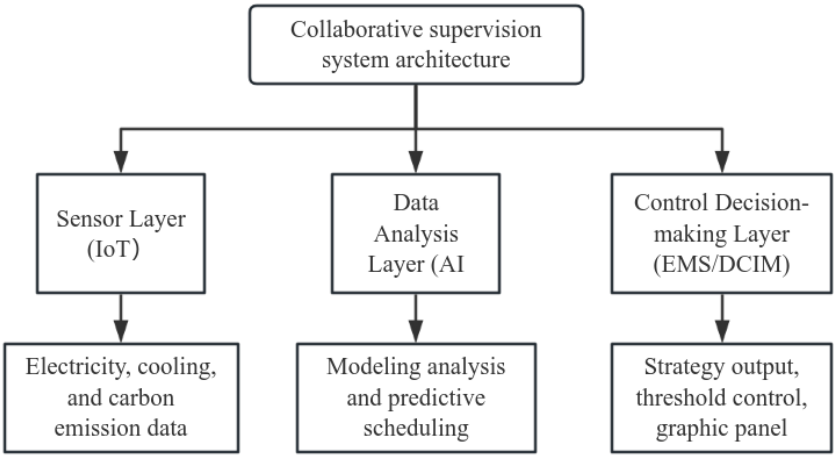


Figure 4. Functional Architecture Diagram of the Energy and environment Collaborative Supervision System for data centers.

5. Conclusion

The low-carbon and sustainable development of data centers is based on the research of three dimensions: the important concept and key technologies of low-carbon data centers, and the development strategy. It achieves energy conservation and efficiency improvement through the application of comprehensive hardware updates, virtualization, and green cooling technologies. Combined with the integrated management of renewable energy and the online operation and maintenance system, it realizes a green and intelligent data center. And strengthen the research on multi-system collaboration and the application of AI technology in the future development trend of data centers to promote the green, low-carbon, modular and autonomous operation and maintenance development of data centers.

References

1. M. Radulescu, J. Cifuentes-Faura, K. Si Mohammed, and H. Alofaysan, "Energy efficiency and environmental regulations for mitigating carbon emissions in Chinese Provinces," *Energy Efficiency*, vol. 17, no. 6, p. 67, 2024. doi: 10.1007/s12053-024-10248-3
2. Q. Wu, and S. Li, "Decarbonization by digits: how data factors drive nonlinear sustainable dynamics in manufacturing," *Applied Energy*, vol. 374, p. 123967, 2024. doi: 10.1016/j.apenergy.2024.123967
3. J. Liu, Y. Zheng, X. Hu, and S. Yu, "Assessing renewable energy efficiency to identify improvement strategies: A network data envelopment analysis approach," *Energy for Sustainable Development*, vol. 76, p. 101308, 2023. doi: 10.1016/j.esd.2023.101308
4. Z. Wang, S. Chen, L. Bai, J. Gao, J. Tao, R. R. Bond, and M. D. Mulvenna, "Reinforcement learning based task scheduling for environmentally sustainable federated cloud computing," *Journal of Cloud Computing*, vol. 12, no. 1, p. 174, 2023. doi: 10.1186/s13677-023-00553-0
5. M. S. Hoosain, B. S. Paul, S. Kass, and S. Ramakrishna, "Tools towards the sustainability and circularity of data centers," *Circular Economy and Sustainability*, vol. 3, no. 1, pp. 173-197, 2023. doi: 10.1007/s43615-022-00191-9

6. V. Jain, and A. Mitra, "Enhancing Role of Innovations in Shaping a Digital Circular Economy," *Sustainable Innovations and Digital Circular Economy*, pp. 249-267, 2025. doi: 10.1007/978-981-96-1064-8_13
7. S. Yadav, P. Kumar, and A. Kumar, "Techno-economic assessment of hybrid renewable energy system with multi energy storage system using HOMER," *Energy*, vol. 297, p. 131231, 2024. doi: 10.1016/j.energy.2024.131231
8. M. Seyedmahmoudian, B. Horan, R. Rahmani, A. Maung Than Oo, and A. Stojcevski, "Efficient photovoltaic system maximum power point tracking using a new technique," *Energies*, vol. 9, no. 3, p. 147, 2016. doi: 10.3390/en9030147
9. N. Phuangpornpitak, and S. Tia, "Opportunities and challenges of integrating renewable energy in smart grid system," *Energy Procedia*, vol. 34, pp. 282-290, 2013. doi: 10.1016/j.egypro.2013.06.756
10. C. J. Lupa, L. J. Ricketts, A. Sweetman, and B. M. Herbert, "The use of commercial and industrial waste in energy recovery systems-A UK preliminary study," *Waste management*, vol. 31, no. 8, pp. 1759-1764, 2011.
11. L. Wang, A. C. Tan, M. E. Cholette, and Y. Gu, "Optimization of wind farm layout with complex land divisions," *Renewable energy*, vol. 105, pp. 30-40, 2017. doi: 10.1016/j.renene.2016.12.025
12. J. C. Cruz, and A. M. Garcia, "Machine Learning for Predictive Maintenance to Enhance Energy Efficiency in Industrial Operations," *ITEJ (Information Technology Engineering Journals)*, vol. 9, no. 1, pp. 15-22, 2024. doi: 10.24235/itej.v9i2.125