

Research Status and Future Development of Cooling Technologies for Green and Energy-Efficient Data Centers

Chen Xintuo¹, Zhou Liyang², Zhang Chengbin³, Wang Shuhua², Zhang Liangliang¹, Chen Jianfeng¹

1. College of Chemical Engineering, Beijing University of Chemical Technology, Beijing 100029, China
2. Juhua Group Corporation, Quzhou 324004, Zhejiang, China
3. School of Energy and Environment, Southeast University, Nanjing 210096, China

Abstract: Data centers are an important infrastructure in China, with constantly increasing industry size, market revenues, and electricity consumption. Major data center cooling systems currently in operation in China include air cooling, which has a low energy utilization rate and significant energy saving potential. Therefore, it is imperative to develop novel cooling technologies for green and energy-efficient data centers. Cooling technology must solve two problems. First, as equipment heat production increases, cooling capacity must be improved to match heat production and transfer rates. Second, to address the low energy efficiency problem, green and energy-efficient cooling technologies with industrial application prospects must be developed. This study reviews data center cooling method development status and categorizes cooling technologies into air conditioner, fresh air, indirect, direct single-phase liquid, and direct two-phase liquid cooling. Cooling technology development trends are analyzed and green and highly efficient data centers are suggested, which focus on improving cooling systems efficiency, optimizing hot and cold fluid channels, applying new liquid cooling materials, and promoting heat recovery systems. Furthermore, we propose recommendations for stable orderly development of the data center industry to optimize top-level design, develop key technologies, formulate scientific standards, improve industrial layout, and focus on energy structure.

Keywords: data center; green and energy saving; cooling method; liquid cooling

1 Introduction

Data are a basic national strategic resource and an important factor in production. Since the 21st century, with the rapid development of information technology (IT) and the rapid increase of applications in the big data era, such as the Internet of Things, the importance of data centers responsible for computing, storage, and data-information interaction has become increasingly apparent. As an important carrier of China's new infrastructure and Internet Plus strategies, data centers are a key cornerstone for China's industrial digitalization, life intelligence, and service informatization transformation, which can support future high-quality economic and social development. The data center industry is valued by countries worldwide, and the market size is increasing annually. Investments, mergers, and acquisitions are active, and competition is fierce. The global data center market exceeded 67.9 billion USD in 2021 and is expected to reach 74.6 billion USD in 2022, with an annual growth rate of approximately 10%. China considers data center industry development significantly important. Large/ultra-large data centers have been developed, and industry revenue has increased. Fig. 1 shows China's data center rack scale and its changes. By the

Received date: May 7, 2022; **revised date:** June 30, 2022

Corresponding author: Chen Jianfeng, professor of Beijing University of Chemical Technology, member of Chinese Academy of Engineering. Major research fields include chemical engineering and materials chemical engineering. E-mail: chenjf@mail.buct.edu.cn

Funding program: National Key R&D Program of China (2021YFB3803200)

Chinese version: Strategic Study of CAE 2022,24(4):94-104

Cited item: Chen Xintuo et al. Research Status and Future Development of Cooling Technologies for Green and Energy-Efficient Data Centers. Strategic Study of CAE, <https://doi.org/10.15302/J-SSCAE-2022.04.010>

end of 2021, this reached 5.2 million racks, with a five-year compound annual growth rate of over 30%. Among the racks, those in large data centers reached 4.2 million, which accounted for 80% of the total. In 2021, China's data center industry revenue reached approximately 150.02 billion CNY, almost three times larger than the market revenue in 2017, with an average annual growth rate of approximately 30% [1]. The continuous advancement of China's industrial modernization has created significant demand for computing requirements, and the market has broad prospects. Traditional and other industries are also actively conducting digital transformation and upgrading, optimizing the industrial structure and generating new demand growth points for the data center industry. Over time and with the development of industrial informatization, the data center industry will be increasingly connected with a growing number of sectors, and is expected to eventually become one of the pillar infrastructure industries.

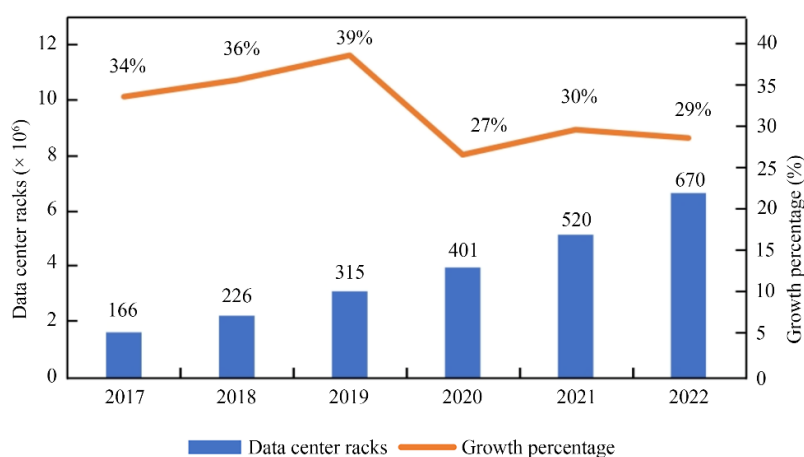


Fig. 1. China's data center rack scale and its changes.

The data center industry in China has entered a new stage of high-speed, high-quality development after experiencing the industry budding and guidance period. With the explosive growth of data center computing traffic capacity, it is increasingly difficult for traditional data center networks to provide low network latency required to support cloud and edge computing, among others. To better meet data processing demands, take full advantage of data center scale benefits, and to reduce business deployment and maintenance costs, large- and ultra-large-scale data centers have become a primary choice for modern data centers. As the scale of data centers expands, computing power and power density increase, and energy consumption supporting IT equipment operation also rapidly increases. The high energy consumption of data computing, storage, and exchange is accompanied by significant heat production in the equipment. IT equipment converts more than 99% of electrical energy into heat energy, and 70% of heat energy must be removed by the data center cooling system [2], which further increases data center power consumption. Fig. 2 shows the Chinese data center electricity consumption over recent years. In the past five years, this has maintained a growth rate of more than 15%. In 2020, power consumption exceeded 200 billion kW·h, accounting for 2.7% of China's total power consumption [3], equivalent to 2×10^8 tons of carbon dioxide emissions (converted to coal-fired power generation). As an auxiliary unit, the power consumption of the cooling system in a large data center is essentially the same as that of the IT equipment. Therefore, the cooling system has a significant energy-saving potential. Data center cooling technology innovation is the key to realizing green and low-carbon development for the entire industry.

This article begins by focusing on the challenges to be urgently solved in terms of the high energy consumption and difficult heat dissipation in data centers. We summarize various developed data center cooling methods that have been applied, and discuss the principles and characteristics of each method. In addition, we present cooling method development trends in the data center industry and propose suggestions for future development goals for green and energy-efficient data centers.

2 Data center cooling technology demand analysis

Data centers must address two major problems: high-energy consumption and difficult heat dissipation. Reliance on the advantages of regional energy supply and the natural environment can reduce the cost of electricity and cooling. Therefore, large data centers can be built in the northwest and southwest regions with sufficient energy and low temperature, which positively exploits the natural environment while reducing operating costs. However, one

problem is that high network latency owing to long data transmission distances makes it difficult to meet edge computing requirements such as high-precision navigation and online control. A realistic requirement remains to build data centers in densely populated and economically developed eastern regions. Under real conditions including a tight energy supply and an unfavorable natural environment, industry consensus has been reached to solve the bottleneck of high energy consumption and difficult heat dissipation and to develop green and energy-efficient data centers. Green and energy-efficient data centers must achieve both heat production and transfer rates for energy efficiency improvement. The development and application of new cooling methods has become a top priority for the energy-efficient and green operation of China's data centers.

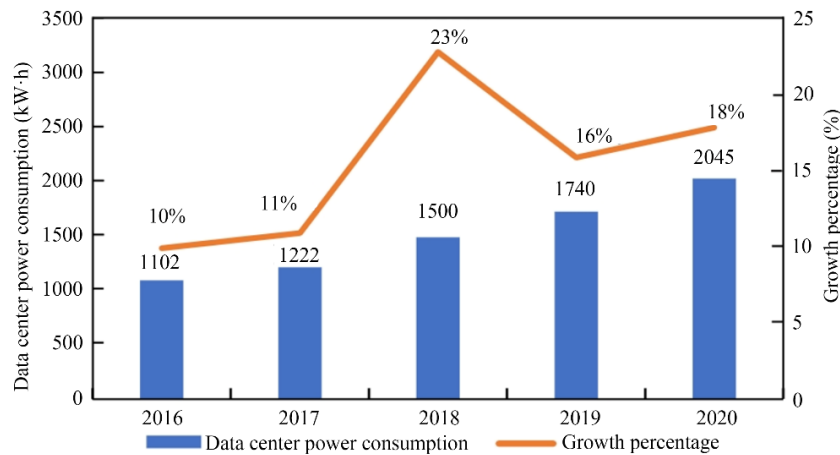


Fig. 2. China's data center power consumption and its changes.

2.1 Matching of heat transfer rate with heat production rate of IT equipment cooling system

Continuous and stable data center operation is the fundamental goal. The key to achieving this is to match heat production and heat transfer rates. As the number of servers in a unit server cabinet increases, the cabinet heat production increases daily, and the requirements for cooling systems continue to increase. The heat production rate can be measured by the power density of the cabinet, which is defined as the energy consumed by a single cabinet for stable operation in kW/r, where r represents a single cabinet. A higher power density of the cabinet means that it produces more heat and a higher heat transfer rate for the cooling system is required. The development of data center server cabinet power from low to high density is inevitable. The low-power-density cabinet is below 5 kW/r, the medium-power-density cabinet is 5–10 kW/r, and the high-power-density cabinet is above 10 kW/r. At present, the power density of data center cabinets currently in use is typically 5–10 kW/r, and that of some super-large data center cabinets in use is above 30 kW/r, even approaching 100 kW/r. This presents a significant challenge for traditional cooling methods.

As shown in Fig. 3 [4], currently developed cooling technologies mainly include air and liquid cooling. Air cooling includes natural and forced air cooling, and the applicable cabinet power density is low. Liquid cooling can be divided into direct single-phase and direct two-phase liquid cooling. The cooling media and methods used in the cooling system are different, and the heat-transfer rate varies significantly. Traditional air cooling can cool cabinets up to 30 kW/r. Using air cooling for cabinets with a power density of more than 30 kW/r, heat production and transfer rates cannot be matched, which causes cabinet temperature to rise continuously, resulting in a decrease in computing power and potential equipment damage. Therefore, for big data, as the power density of cabinets is rapidly increasing, continuous innovation in cooling systems and methods is required to improve the heat transfer rate.

2.2 Low energy efficiency that requires green high-efficiency cooling methods

Heat dissipation consumes energy, therefore, the increase in heat transfer rate must consider energy efficiency. The indicator used to measure the overall energy consumption level of a data center is power usage effectiveness (PUE), which is defined as the ratio of total data center energy consumption to IT equipment energy consumption. Under the same IT power consumption, a PUE closer to 1 indicates that the data center has lower non-IT power consumption and the higher energy efficiency than other centers. As shown in Fig. 4, the total energy consumption of a data center consists of power supply and distribution, lighting, cooling, and IT equipment. As shown in Fig. 5, approximately 43% of data center energy consumption in China is currently used for heat dissipation and cooling

(The PUE of the corresponding data center is larger than 2), which has a high cooling cost and significant energy-saving potential [5].

To improve the data center energy efficiency level in China and promote green development in the information industry, the Ministry of Industry and Information Technology of the People’s Republic China issued the *Three-Year Action Plan for New Data Center Development (2021–2023)*, requiring that the PUE of newly built large-scale and above data centers must be under 1.3. The National Development and Reform Commission issued *Opinions on Reinforcing Energy Efficiency Constraints to Promote Energy Conservation and Carbon Reduction in Key Areas* that clearly encourages key industries to use new infrastructure such as green data centers to achieve energy conservation and consumption reduction. Taking the energy efficiency level of advanced green data centers domestic and international as the target of technological transformation, data center PUE will not exceed 1.5 in general by 2025.

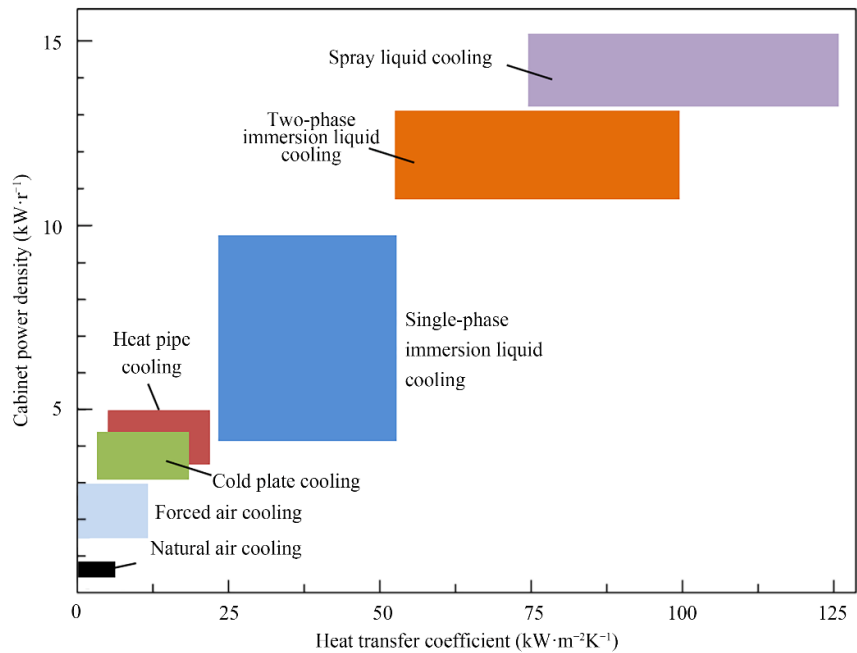


Fig. 3. Cooling type diagram.

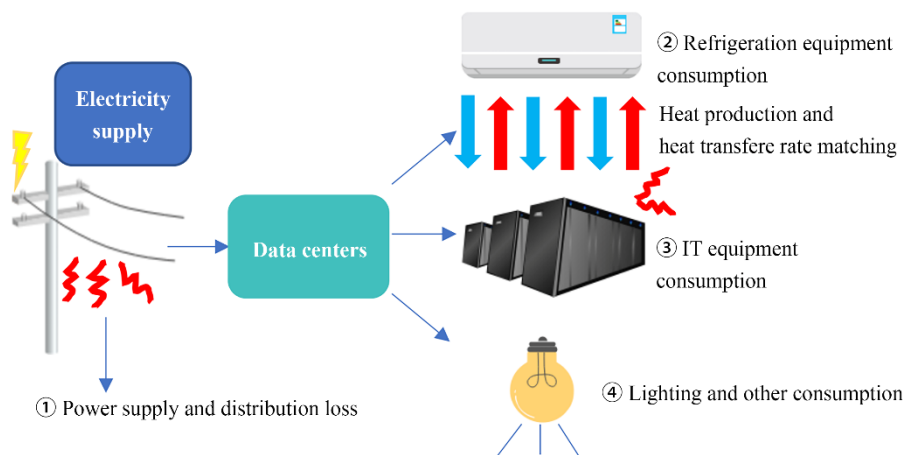


Fig. 4. Schematic of data center energy consumption composition.

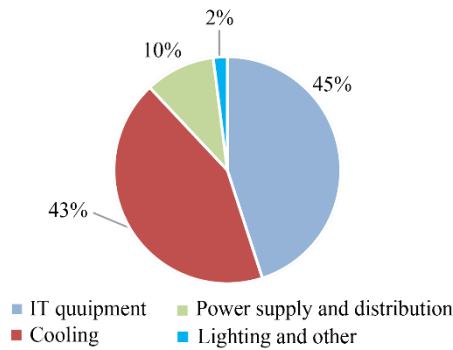


Fig. 5. Schematic of energy consumption distribution of data centers in China.

3 Air-cooling technology development status

Air cooling is a method that delivers cold air to IT equipment for heat exchange, which played a significant role in the initial IT industry stage. At that point, the concept of a data center was not clarified, and was generally known as a computer room or center. The computer center consisted of a few businesses, small scale, and low data processing capacity with the cabinet power density at approximately 1–2 kW/r. Air cooling was the most suitable cooling method for small and micro data centers during that period [6]. Once large data centers with high power densities gradually become mainstream, the air-cooling method could continue to meet the heat dissipation requirements of servers with cabinet power densities below 30 kW/r. However, its energy efficiency is generally low and it does not meet the green development requirements of the IT industry [7].

According to the air flow source, air cooling can be divided into natural and forced air cooling. The air near the IT equipment is heated, its temperature increases, and density decreases. The hot air flows toward the space above the data center, creating a pressure differential that allows the cool air to flow naturally toward the IT equipment. Natural air cooling has low air velocity, insignificant convective heat transfer, poor cooling effect, and low energy efficiency. Therefore, this cooling method was eliminated.

Forced air cooling is a method that uses draught fans to force cold air to flow into the IT equipment. It is currently the most widely used cooling method for data centers. This method achieves good cooling performance by increasing the heat dissipation area of the IT equipment or accelerating the air flow rate. Studies have demonstrated that the use of technologies such as reasonable airflow organization, intelligent ventilation, precise air supply, intelligent heat exchange, air-conditioner cycling, polarized oil additives, and air conditioning compressor refrigerant control can reduce the energy consumption of forced air-cooling data centers by approximately 20% [8]. According to different air sources, forced air cooling mainly includes two methods: air-conditioner and fresh air cooling.

3.1 Air-conditioner cooling

Air-conditioner cooling is a method in which air conditioners cool the air inside data centers. Owing to the high energy consumption of air conditioning systems, data center PUE using this cooling method is generally above 1.7 [9]. Air-conditioner cooling has good adaptability, straightforward maintenance, mature technology, and high stability. However, its cooling efficiency is low, and it is only suitable for small- and medium-power-density cabinets. Hence, this option cannot address the cooling requirements of medium and large data centers in the future.

3.2 Fresh air cooling

Fresh air cooling refers to a cooling method that uses the outside air (fresh air) of the data center as the cooling medium. When the internal temperature of the data center is higher than that of the external environment, external cold air can be directly transported into the data center through draught fans and the heated air can subsequently be discharged outside through the channel. This method can save 40% of the energy compared with air-conditioner cooling; however, it is sensitive to air quality conditions; therefore, it is necessary to add dust filter and dehumidification systems to control the air quality. Yahoo's data center in Brockport, New York, and Facebook's data center in Prineville, Oregon use fresh air cooling for heat dissipation, and the PUE can be reduced to below 1.2. The current fresh air cooling technology is developing rapidly, such as the design of Distributed Airflow Cooling System and the Blowing and Drawing Ventilation Cooling System, which can realize data center temperature control without air conditioners throughout the year [10]. Fresh air cooling data centers have high requirements for

environments. In temperate or frigid zones with low average temperatures throughout the year, the use of fresh air cooling can reduce cooling costs.

4 Liquid cooling technology development status

Air cooling methods use air as the cooling medium, which has low thermal conductivity. The thermal conductivities of liquids are an order of magnitude higher than that of air. In theory, liquid cooling can significantly improve the heat transfer rate and meet the cooling requirements of high-power-density cabinets. Considering China's data center market revenue in 2021, the liquid cooling data center market revenue reached 45–50 billion CNY, accounting for more than 30%, and significant room for development remains.

A certain research foundation has been laid with progress made in data center liquid cooling both domestically and internationally, which is becoming a transformative technology. Liquid cooling methods can be divided into three categories: indirect, direct single-phase, and direct two-phase liquid cooling.

4.1 Indirect liquid cooling

Indirect liquid cooling is a method in which the liquid medium is in indirect contact with the heat-generating components using good conductors of heat. The liquid undergoes a phase change or non-phase change in the fluid channel to absorb heat, and the heat-generating components cool down. In a typical indirect liquid-cooling direct-current system, traditional air-cooling fins are replaced by evaporators or liquid-cooling fins [11], and a cooling medium distribution and delivery system is added. The cooling medium is connected to the IT equipment from outside the data center through fluid pipelines. It then exchanges heat with an external cooling source, cools down, and is transported to the inside of the data center through pipelines for cooling circulating. Because electronic chips are the main source of heat, the cooling medium typically cools electronic chips by indirect cooling. Remaining components, such as hard drives, are cooled by air [12]. Currently, except for electronic chips, no indirect cooling schemes exist that can be implemented commercially for other electronic components [13].

In indirect liquid cooling, because the liquid is not in direct contact with the IT equipment, a wide range of choices regarding the liquid medium exist. It is only necessary to consider whether the thermal conductivity meets technical requirements and whether it is compatible with the fluid pipelines. Indirect liquid cooling can be divided into cold-plate and heat-pipe cooling. Currently, most liquid media do not undergo phase change during cold-plate cooling, whereas the liquid phase change occurs during cold-plate cooling.

4.1.1 Cold-plate cooling

Cold plate cooling is a method which is used at the chip level. As shown in Fig. 6, cold-plate cooling attaches metal cold plates to electronic chips, and liquid flows in the cold plates. When the chips heat up, they conduct heat to the cold plate. When the liquid flows through cold plates, it heats up, uses its sensible heat to remove heat from chips, and exchanges heat with the external cold source through pipelines. Water is the most widely used cooling medium for this purpose. Cold-plate cooling is the most widely used cooling method for liquid-cooling data centers. It uses a combination of liquid and air cooling. Liquid cooling is used for chips whereas air cooling is used for other electrical components, such as hard drives. However, this is not strictly liquid cooling [14]. Compared with air cooling, which can cool cabinets up to 30 kW/r, the cold plate can cool cabinets to less than 45 kW/r, which is more energy-efficient and less noisy. It does not require expensive water chilling units and has certain advantages over pure liquid cooling [15]. Tencent used a combination of cold plates and IT equipment micro-modules to create a liquid cooling micro-module data center. The PUE of the liquid cooling module is close to 1.1, and the overall PUE is 1.35. In 2019, Microsoft submerged a data center capsule containing 864 servers into the sea. Seawater was pumped into pipelines in a sealed cabin to cool the servers through indirect heat exchange, and a cooling performance test was conducted. Based on their experiment, it was identified that the management and construction costs of the subsea data center were better than those of the traditional data center, and the failure rate of IT equipment in water was one-eighth of that on land.

4.1.2 Heat-pipe cooling

Heat-pipe cooling is a method that adds heat pipe elements into IT equipment, uses the rapid heat transfer properties of heat conduction and liquid phase change, and quickly transfers equipment heat to the outside. The liquid absorbs heat and changes phase at the end of the heat pipe, which is in contact with the heating component, and uses the latent heat of the phase transformation to absorb heat and convert it into a gaseous state. The gaseous medium flows to the condensation section based on the pressure difference, and a condensation phase change occurs

to release heat at the condensation section. The condensate flows back to the evaporation section through the capillary or gravity action of the heat pipe wick, and circulates and transfers the heat of the heated device. The condensing section is then cooled, and heat is exchanged by air or liquid cooling [16]. The thermal conductivity of heat pipes exceeds that of all known metals, and the cooling medium includes water, methanol, acetone, ammonia, dichlorofluoromethane, and hydrated silica. The condensing section can be cooled using water or air [17]. The heat pipe is not connected to the cooling medium connection pipelines of the chips, which reduces the risk of cooling medium leakage inside the IT equipment. In addition, a pump is not required, and the capillary pressure difference is used to drive the two-phase cooling medium self-circulation movement to generate heat exchange. Because there are no moving parts in the cooling structure, reliability of heat-pipe cooling systems is high [18]. The heat-pipe cooling cabinet does not require reservation for air cooling convection space, which improves rack power density while achieving a good cooling effect, efficient equipment heat dissipation, and waste heat recovery and reuse [19]. Based on loop heat pipe technology, Tsinghua University applies a multi-stage loop heat pipe to data centers. A loop heat pipe cabinet developed and transformed into a data center in Beijing could save energy by 41.6% compared with air conditioning and refrigeration [20]. Because the cooling capacity of heat pipes is not significantly improved compared to that of cold plates, it is not currently considered for data center cooling.

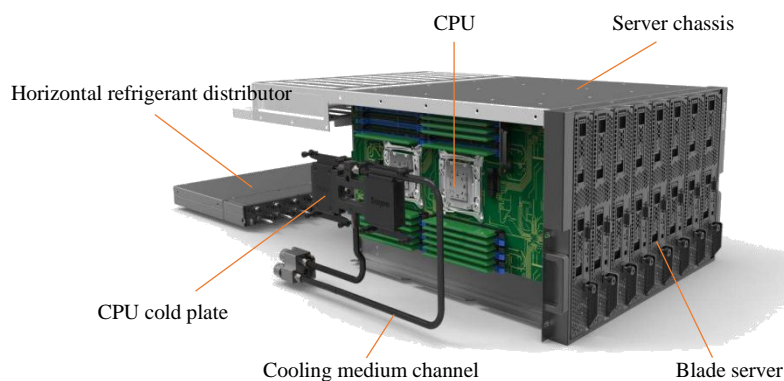


Fig. 6. Cold-plate cooling.

4.2 Direct single-phase liquid cooling

Direct single-phase liquid cooling is a method in which an insulating liquid that does not affect the normal operation of IT equipment components, is in direct contact with the components, and the liquid does not undergo a phase change to remove heat.

4.2.1 Single-phase immersion liquid cooling

Single-phase immersion liquid cooling is a method in which IT equipment is immersed in a sealed tank containing a cooling medium. Moving parts in the equipment are not required and only the fluid flow channel must be planned in advance. No phase change occurs when the cooling medium exchanges heat in the equipment. The cooling medium uses sensible heat to remove heat. The heated liquid flows from the tank using the circulating pump, enters the cooler to cool down, and flows back to the tank to circulate heat exchange. The principle of single-phase immersion liquid cooling is relatively straightforward and can be applied to cabinets up to 100 kW/r. As shown in Fig. 7, the technical difficulties associated with this method include the screening of liquid cooling materials and design of the corresponding IT equipment.

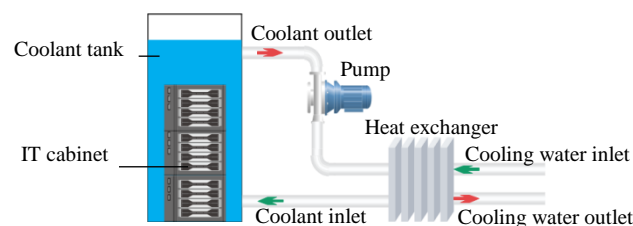


Fig. 7. Schematic of single-phase immersion liquid cooling.

For direct single-phase liquid cooling materials, it is necessary to satisfy the performance requirements of high insulating properties, low viscosity, high flash point or non-combustibility, low corrosiveness, high thermal stability, and low biological toxicity. The liquid cooling material can be mineral oil or fluoride, such as perfluoro amine or perfluoropolyether. For example, the “Juxin” direct single-phase liquid cooling material from the Juhua Group Corporation has entered the stage of industrialization, and its main performance indicators are basically the same as those of foreign monopoly products. The design of the equipment is different from that of air cooling. It is necessary to reserve liquid flow channels inside IT equipment, which cannot contain chemical compositions that interact with the materials. Immersion liquid cooling cabinets do not have sealed enclosures and can effectively protect equipment from dust or sulfide, allowing low-maintenance hot swapping without power outages. [12] However, additional maintenance procedures are required to prevent the loss of the cooling medium and to remove infiltrated air and moisture [21]. As shown in Fig. 8, Alibaba Cloud Computing Co., Ltd. has developed the world’s largest single-phase immersion liquid cooling data center (Renhe Data Center) in Hangzhou using self-developed liquid cooling server clusters. This is the first liquid cooling data center in China with a green rating of 5A, and a PUE as low as 1.09. During data center operation, engineers selected thousands of liquid-cooling servers for comparison with air-cooling servers. Based on the failure rate of each server component, for liquid cooling servers, this decreased by 53% [22]. Immersion liquid cooling maintains temperature below the heat production limit, which effectively improves chip efficiency.



Fig. 8. Alibaba Cloud immersion liquid cooling server.

4.2.2 Single-phase spray liquid cooling

Single-phase spraying uses liquid nozzles to spray the cooling medium on the surface of the heating component; the droplet size is large, and no phase change occurs. The liquid forms a thin boundary layer on the surface of the component for heat exchange, which can locally produce a strong single-phase convective heat transfer effect and use sensible heat for dissipation [23]. Prior to coming into contact with the heat-source surface, the cooling medium was atomized or dispersed into small droplets. This process is mainly performed by the pressure difference of the liquid nozzle [24]. Guangdong Heyi New Materials Research Institute Co., Ltd. developed a centralized supply spray liquid cooling server. Vegetable oil is sprayed on the chips, without phase change during the entire process. The PUE is less than approximately 1.1.

Once spray cooling increases nozzle pressure, the cooling medium is ejected at high speed, which can cause jet impingement. After the cooling medium is sprayed directly onto the heated surface, it hits the surface at high speed. Using the principle of high-speed forced convection heat transfer and a short stroke, this forms a thin temperature and velocity boundary layer and produces a strong heat transfer effect on the impact area. Combined with bubble-induced flow mixing and latent heat transfer, high heat transfer efficiency can be achieved [25].

A single-phase spray is a chip-level heat exchange with high cooling efficiency. It is technically difficult to set spray nozzles on each heat-generating component, and this has not yet been applied to data center cooling systems. When the liquid contacts the high-temperature heat source during the spraying process, floating and evaporation phenomena occur, and mist droplets and gases are emitted to the outside of the chassis along the holes and gaps, resulting in decreased cleanliness in the equipment room environment with the potential to impact other IT equipment. Therefore, spray liquid cooling has a high requirement for sealing the chassis.

Spray cooling can also be used in combination with cold-plate cooling to spray liquid onto cold plates for indirect heat exchange to replace air cooling inside IT equipment. The technical difficulty of this method is low and it has been currently applied. SparyCool designed a spray indirect heat exchange system for the US Army Command Post Platform, which improved the convenience and disassembly of the vehicle command platform. In this manner, the equipment is packaged in a separate shipping case as a functional network configuration unit for easy mobile and

static operations. It only requires a power supply and network connection, without reconfiguration and rewiring of network equipment and hardware, and is suitable for mobile operations [26].

4.3 Direct two-phase liquid cooling

Direct two-phase liquid cooling has one key innovation, that is, to use the latent heat of phase change of media for cooling, with boiling points below 80 °C, in general.

4.3.1 Two-phase immersion liquid cooling

As shown in Fig. 9, the two-phase immersion liquid comes into contact with the heat source in the cooling-medium tank. Vapor bubbles form on the surface of the heat source and rise to the area above the tank, and the vapor converts back to liquid in the condenser after the heat exchange and flows back into the tank. Phase-change heat transfer involves latent heat transfer, gravity-driven two-phase flow, and bubble-induced flow mixing [27]. In addition, the phase change process is conducted on the heated parts of the IT equipment, and the formation of bubbles affects the subsequent heat transfer of the cooling medium, which is technically difficult and requires high-quality sealing of the cooling medium tank. To meet the requirements of direct single-phase liquid cooling, two-phase immersion liquid cooling materials require a lower boiling point, which uses the latent heat of phase change to absorb heat within the stable operation range of IT equipment. It can use short-chain fluorides, such as FC-72, Novec-649, HFE-7100, or PF-5060, which can sustain chip temperature below 85 °C and ensure chips run stably [12]. The steam can be returned to the tank by gravity after being cooled in the condenser without the necessity to provide additional circulating power, and the power density of a single cabinet can reach more than 110 kW/r. As shown in Fig. 10, Dawning Information Industry Co., Ltd. released the “Silicon Cube” high-performance computer in 2019. This computer uses immersed phase-change liquid cooling, which is considered a breakthrough for reaching the energy efficiency level limit of a single computer, where the PUE reaches 1.04. This demonstrates the advantages of liquid cooling technology in high-performance computers; however, its application in computer clusters, such as data centers, remains to be explored.

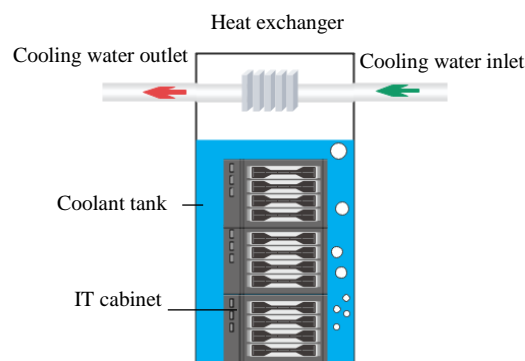


Fig. 9. Schematic of two-phase immersion liquid cooling.

4.3.2 Two-phase spray liquid cooling

Two-phase spraying is a cooling method that atomizes the cooling medium into tiny droplets to form a liquid film on the surface of heating components on the basis of single-phase spraying, and the liquid does not flow significantly and is converted into vapor to remove heat [28]. Its cooling efficiency is more than three times that of single-phase spray liquid cooling [29]. The heat-transfer mechanism of two-phase spray liquid cooling is complex. Mechanisms such as thin-film evaporation, single-phase convection, and secondary nucleation have been proposed by researchers; however, no unified heat transfer model exists, which must be further researched [30]. Two-phase spray cooling is the method known to consist of the strongest heat-removal capability. It has broad prospects albeit high energy consumption and is suitable for cooling cabinets with a power density greater than 140 kW/r. However, no commercialized applications have been developed for data centers. The stability of the two-phase spray system must be verified. The nozzle is easily blocked and its maintenance is difficult. The system requires a sealed vapor chamber and a vapor recovery device, and issues such as the reliability of electronic equipment must also be considered [27, 30].



Fig. 10. Silicon cube.

5 Development trends for cooling technology in green and energy-efficient data centers

Data centers are Internet infrastructure, and its transformation and upgrading plays an important role in China's social modernization. The development direction of cooling technologies for green and energy-efficient data centers predominantly includes improvement in cooling systems efficiency, optimization of hot and cold fluid channels, application of new liquid cooling materials, and promotion of heat recovery systems.

5.1 Cooling systems efficiency improvement

The cooling system of a data center includes cooling towers, water-chilling units, pumps, condensers, fluid pipelines, and other equipment. Reasonable selection of process conditions and equipment can significantly improve overall efficiency, thereby improving the energy efficiency ratio. Energy efficiency can be improved using efficient uninterruptible power supplies, flexibly changing the frequency of pumps and chillers with load drops or seasonal changes, and optimizing the internal layout of the data center [31].

5.2 Optimization of hot and cold fluid channels

Hot and cold fluid channels are also known as hot aisle cold aisle (HACA) technology and were first applied in air cooling data centers. The entrances or exits of adjacent racks are opposite, forming separate hot and cold aisles. Cold air input to the server by the air-conditioning system is separated from the hot air after the heat exchange with IT equipment, and the cold and hot aisles are input to or output from the server, respectively [32], which improves the heat exchange efficiency of the cooling medium from the perspective of thermodynamics. The cold air is input into the server through the cold aisle, and the hot air after heat exchange with the IT equipment is output through the hot aisle [32]. Cold and hot air are not mixed, which improves heat exchange efficiency. Ensuring airtightness of the hot and cold aisles also assists in reducing the air conditioner fan power by approximately 15% and saves energy [33]. HACA technology can also be used as a reference for future liquid cooling data centers. By optimizing the internal structure of IT equipment, fluid channels for cooling medium input and output are formed to avoid mixing of cooling media with different temperatures. Thus heat-exchange efficiency is thermodynamically optimized. Second, it can also increase the turbulence degree of the cooling medium, which thins the thermal boundary layer and reduces convection thermal resistance.

5.3 New liquid cooling material applications

Unlike air-cooling data centers, which use air as the cooling medium, liquid-cooling data centers have multiple types of cooling media and an abundance of choice. The identification of liquid cooling materials that are compatible with IT equipment, with a high degree of safety, and excellent heat transfer performance represent critical research directions for future cooling technology development. Liquid cooling materials selection requires comprehensive consideration of various factors, such as flash point, volatility, biological toxicity, thermal conductivity coefficient, viscosity, dielectric constant, and corrosivity. The reason for the difference in the macroscopic properties of liquid cooling materials is the molecular structure. Screening liquid-cooling materials with different molecular structures through simulation and other means, establishing and improving a liquid-cooling material database, and identifying competitive low-cost high-performance liquid-cooling materials represent the cornerstones for data center transformation and upgrading from air- and semi-air cooling to liquid cooling.

5.4 Heat recovery systems promotion

Most current heat recovery technologies capture waste heat by exchanging heat with a cooling-fluid stream. The quality and quantity of waste heat are determined by the type and size of the thermal management system. In an efficient air-cooling data center, cold air is typically supplied at 25 °C and can be increased to 40 °C after heat exchange with IT equipment [34]. Hot air can convert heat energy into electric power through absorption refrigeration equipment to realize the cogeneration of heat and power. The electricity produced can be used for cooling water, thus to save electric power [35]. After using liquid cooling methods, such as cold plates, an inlet water temperature of approximately 60 °C can maintain the working temperature of IT equipment chips at 85 °C. The heat energy produced by the cooling medium after heat exchange has a higher quality and wider range of use [36]. The recovered waste heat can be used for plant or district heating/hot water production, absorption cooling, organic Rankine cycle, piezoelectric, thermoelectric, biomass fuel drying, desalination/clean water production, among others [37]. By analyzing different types of data center cooling technologies and matching corresponding heat recovery systems, energy efficiency can be effectively improved.

6 Data center cooling technology development vision

The data center is a hub and application carrier of data. It is not only an important part of modern infrastructure, but also the basis for the development of new IT infrastructure and an important precondition for building information platforms, which plays an underlying supportive role for advancing the digital economy. However, owing to complex problems including high energy consumption and difficult heat dissipation, green and low-carbon development of data centers must address significant challenges. With data center development trends becoming larger and denser, cooling technologies must adapt to changes and transformations, and upgrade in a timely manner. It is necessary to eliminate old data centers with serious energy waste and build a new generation of green and energy-saving data centers using liquid cooling. Realizing the transformation and upgrading of data center cooling technologies could assist in supporting China's 14th Five-Year Plan and promote clean low-carbon IT industries. Transformation could also upgrade the industrial base and modernizing the industrial chain, transform the economic development from an extensive to an intensive type, and achieve breakthroughs for carbon peaking and carbon neutrality technologies in the digital economy. The transformation can also ensure that the development of strategic emerging industries such as the digital economy is not restricted by infrastructure. Conducting scientific and technological innovations to identify new economic growth points and appropriate driving forces is of great significance to China's economic growth and social stability.

6.1 Optimizing the top-level design of the regional layout of data centers, and developing high-efficiency cooling technologies

For existing data centers, the cost of transforming from original cooling forms to liquid cooling is very high; therefore, building new liquid cooling data centers on a well-planned basis has greater economic advantages. Data center service life is approximately ten years, and cooling technology must be planned in advance to maintain a steady upward development trend. It is necessary to strengthen national data center management, optimize top-level design and overall coordination, advance the layout, and conduct unified construction. Construction requirements for multiple small data centers could be consolidated into large data centers with high power density, to reduce the cost of liquid cooling and improve energy efficiency. Data centers throughout China should be planned perceiving the situation as a whole with consideration for the unique advantages of local land and energy technology, among other aspects. To effectively achieve a reduction in research and development (R&D) cost and construction of liquid cooling data centers, reasonable industrial layout, intensive resource allocation, and highly centralized organizational structures are required.

6.2 Developing key breakthrough technologies for full industrial chain implementation

China's liquid cooling data centers are in the early stage of development. Factors that affect widespread adoption and promotion of liquid cooling servers include the incomplete industry chain, high cost of equipment procurement and cooling media, few procurement channels, and low compatibility of electronic components [38]. Development of a green data center must be a systematic project from planning to implementation involving a series of technical challenges to be addressed. Multidisciplinary joint research is required from the screening and preparation of liquid cooling materials to the optimization of power supply and distribution systems to reduce loss, including data center security control systems upgrades. It is necessary to increase investment in scientific research, promote all key links

from cooling materials to adapting servers simultaneously, and organize regional science and technology innovation complexes. It is important to focus on breakthroughs in common, key and stranglehold technologies, address industrial bottlenecks, and reduce R&D costs. In addition, innovation and development initiatives should be firmly grasped such that green data center equipment is capable of achieving independent R&D, production, and construction, finally achieving world leadership. Made in China should be transformed into Created in China to enhance the core competitiveness of the data center industry.

6.3 Formulating scientific standards to guide the healthy development of the data center cooling technology industry

Green data center industry standards should be developed to formulate unified specifications for IT equipment, liquid cooling materials, data center operation and maintenance, power supply and distribution, safety, and thermal control protection. A scientific organic system should be formed to improve efficiency from R&D to production. It is necessary to ensure smooth development of modern heat dissipation and cooling materials, IT equipment, and other R&D aspects. Key technologies in different industries should be rationally integrated and timely transformed into productivity. Deep integration between different technologies and industries is considerably beneficial. Realizing industrial scientific management, cross-border integration demonstration projects, and reducing operation and maintenance costs to promote effective development of the data center market, should be prioritized. It is anticipated that the future data center industry will create new application scenarios coupled with new consumer demands.

6.4 Improving industrial layout through coordinated promotion of data centers

The geographical distribution of data centers determines the most appropriate cooling technology application scheme, and coordinated development of air and liquid cooling technology should be realized in China. Central and western regions have abundant electricity, low electricity and land prices, low average temperatures throughout the year, and clean air, which meets the requirements of a fresh air cooling environment. This is suitable for the construction of fresh air cooling data center zones in the central and western regions, which is in line with the development trend of green energy-efficient data centers. However, owing to low thermal conductivity of the air, it cannot cool high-power-density servers. From the perspective of cooling costs, high-availability and high-density server platforms, such as blades, should be used as little as possible in fresh air cooling data centers. For industries with special requirements for network delay and server type, such as security transactions, high-tech, financial, or communication, it is not recommended to conduct data processing through fresh air data centers in the central and western regions to avoid long-distance data transmission. The establishment of data centers in large- and medium-sized cities in east China requires a comprehensive consideration of many factors. First, land resources are scarcely available and costly. Second, the population is dense, and noise, such as that of air-conditioning units, has a significant impact on the normal production and life of residents. Finally, high operation and maintenance costs exist, such as electricity prices. Considering such factors, a liquid-cooled data center with high power density, quiet operation, and flexible location is more suitable for the eastern region.

6.5 Focusing on clean energy and waste heat utilization

Data centers support significant amounts of data storage, and computing is essentially an energy-intensive industry. To achieve carbon reduction and neutrality targets for data centers, it is necessary to optimize their energy supply structure. Renewable energy forms, such as wind energy and photovoltaic power generation, can be fully integrated into the data center industry. Although this type of energy is intermittent and unstable, and the power structure of a data center can be improved through scientific design to increase the proportion of green electricity. Simultaneously, the price and availability of renewable energy is relatively flat, making data centers less affected by energy supply. In addition, a data center generates a large amount of low-grade heat while consuming electricity. The waste heat production by the data center could be considered for secondary use, such as providing winter heat sources for office spaces or for a combination with absorption cooling technology to convert waste heat into cold energy for recycling.

References

- [1] China Academy of Information and Communications Technology (CAICT). Data center white paper (2022) [R]. Beijing: China Academy of Information and Communications Technology (CAICT), 2022. Chinese.

- [2] Lu T, Lyu X S, Remes M, et al. Investigation of air management and energy performance in a data center in Finland: Case study [J]. *Energy & Buildings*, 2011, 43(12): 3360–3372.
- [3] Ji X L. Ecological and green development: The installed capacity of power generation is getting cleaner and the data center is getting more and more energy-efficient [N/OL]. *China Economic Herald*, (2021-07-29) [2022-05-02]. <http://www.chinadevelopment.com.cn/news/ny/2021/07/1737167.shtml>. Chinese.
- [4] Bao K L, Wang X H, Fang Y B, et al. Effects of the surfactant solution on the performance of the pulsating heat pipe [J]. *Applied Thermal Engineering*, 2020, 178: 115678.
- [5] CCID. White Paper on the Development of China Liquid-Cooled Data Centers [R/OL]. Beijing: CCID Consulting Co., Ltd., (202012) [2022-04-15]. <http://www.ysdj.cn/Private/Files/20210107/6374561177902446434712543.pdf>. Chinese.
- [6] Zhu Y Z. The application and development of data center refrigeration technology [J]. *Standardization of Engineering Construction*, 2015 (8): 62–66. Chinese.
- [7] Jing C G, Li N, Song X W. Effect of different power density racks on data center [J]. *Telecom Power Technologies*, 2019, 36(10): 4. Chinese.
- [8] CHENG B, WANG T, WU H, et al. The strategy and application of energy saving and emission reduction in China telecom data center [J]. *Energy Conservation*, 2012, 31(1): 4-8. Chinese.
- [9] Garimella S V, Yeh L T, Persoons T. Thermal management challenges in telecommunication systems and data centers [J]. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 2012, 2(8): 1307–1316.
- [10] Liu Y, Guan R B, Jing M, et al. Research of data center fresh air ventilation cooling system [J]. *Lecture Notes in Electrical Engineering*, 2014, 262: 299–306.
- [11] Geng H Y. *Data center handbook* [M]. Hoboken: John Wiley & Sons, 2014.
- [12] Tuma P E. The merits of open bath immersion cooling of datacom equipment [C]. Santa Clara: Proceedings of the 2010 26th Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM), 2010.
- [13] ELLSWORTH M J. Flow network analysis of the IBM Power 775 supercomputer Water Cooling System [C]. Orlando: Proceedings of the 2014 IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm), 2014.
- [14] Xiao X W. On heat recovery in data centers cooled by cold plate direct contact liquid cooling system [J]. *Building Energy Efficiency*, 2020, 48(2): 69–73. Chinese.
- [15] Xie L N, Guo L. Discussion on liquid cooling technology and development [J]. *Information and Communications Technology and Policy*, 2019, 296(2): 29–32. Chinese.
- [16] Han X H, Wang X H, Zheng H C, et al. Review of the development of pulsating heat pipe for heat dissipation [J]. *Renewable and Sustainable Energy Reviews*, 2016, 59: 692–709.
- [17] Kheirabadi A C, Groulx D. Cooling of server electronics: A design review of existing technology [J]. *Applied Thermal Engineering*, 2016, 105: 622–638.
- [18] Mcglen R J, Jachuck R, Lin S. Integrated thermal management techniques for high power electronic devices [J]. *Applied Thermal Engineering*, 2004, 24(8–9): 1143–1156.
- [19] Zheng S M. Application of heat pipe technology in data center cooling system [J]. *Journal of Green Science and Technology*, 2019 (24): 187–188. Chinese.
- [20] Tian H, Li Z. Distributed cooling scheme of data center based on loop heat pipe technology and its application [J]. *World Telecommunications*, 2011 (10): 48–52. Chinese.
- [21] EL-GENK M S. Immersion cooling nucleate boiling of high power computer chips [J]. *Energy Conversion & Management*, 2012, 53(1): 205–218.
- [22] Zhong Y F, Guo R, Zhang J, et al. Long-term reliability evaluation of single-phase immersion liquid-cooling server based on electronic fluoride [J]. *China Telecommunications Trade*, 2021 (S1): 55–60. Chinese.
- [23] Park J E, Thome J R. Critical heat flux in multi-microchannel copper elements with low pressure refrigerants [J]. *International Journal of Heat & Mass Transfer*, 2010, 53(1–3): 110–122.
- [24] Kim J. Spray cooling heat transfer: The state of the art [J]. *International Journal of Heat & Fluid Flow*, 2007, 28(4): 753–767.
- [25] Guo D, Wei J J, Zhang Y H. Enhanced flow boiling heat transfer with jet impingement on micro-pin-finned surfaces [J]. *Applied Thermal Engineering*, 2011, 31(11–12): 2042–2051.
- [26] Johnston A, Stone D, Cader T. SprayCool command post platform for harsh military environments [C]. San Jose: 2008 Twentyfourth Annual IEEE Semiconductor Thermal Measurement and Management Symposium, 2008.
- [27] Qiu L, Dubey S, Choo F H, et al. Recent developments of jet impingement nucleate boiling [J]. *International Journal of Heat and Mass Transfer*, 2015, 89: 42–58.
- [28] Liu S Y, Ruan L, Kang Y H. Experimental investigation and numerical analysis of atomization characteristics at the spray cooling axis [J]. *Science Technology and Engineering*, 2018 (2): 225–229. Chinese.
- [29] Yang J, Chow L C, Paris M R. Nucleate boiling heat transfer in spray cooling [J]. *Journal Heat Transfer*, 1996, 118(3): 668–671.

- [30] Silk E A, Gollhofer E L, Selvam R P. Spray cooling heat transfer: Technology overview and assessment of future challenges for micro-gravity application [J]. *Energy Conversion & Management*, 2008, 49(3): 453–468.
- [31] Rasmussen N. An improved architecture for high-efficiency, high-density data centers [J]. *Green Grid White Paper*, 2008, 126: 1–25.
- [32] Joshi Y, Kumar P. *Energy efficient thermal management of data centers* [M]. Boston: Springer, 2012.
- [33] Takahashi M, Uekusa T, Kishita M, et al. Aisle-capping method for airflow design in data centers [C]. San Diego: proceedings of the Telecommunications Energy Conference, 2008.
- [34] Bash C E, Patel C D, Sharma R K. Efficient thermal management of data centers—Immediate and long-term research needs [J]. *Hvac & R Research*, 2003, 9(2): 137–152.
- [35] Ayoub D S, Bruno J C, Saravanan R, et al. An overview of combined absorption power and cooling cycles [J]. *Renewable & Sustainable Energy Reviews*, 2013, 21: 728–748.
- [36] Brunschwiler T, Smith B, Ruetsche E, et al. Toward zero-emission data centers through direct reuse of thermal energy [J]. *IBM Journal of Research and Development*, 2009, 53(3):1–13.
- [37] Ebrahimi K, Jones G F, Fleischer A S. A review of data center cooling technology, operating conditions and the corresponding low-grade waste heat recovery opportunities [J]. *Renewable & Sustainable Energy Reviews*, 2014, 31(2): 622–638.
- [38] Sun C F, Zhang L. Development Trend and Application of Data Center Server Technology [J]. *Communications World*, 2021 (10): 37–38. Chinese.