



OPEN Optimising thermal performance in data centre server racks via a parametric layout configuration study

Xinkun Gao¹, Mingqin Cai¹, Ruiyong Mao¹ , Cheng Hu¹ , Hongwei Wu², Qiaoxinyi Zhou¹, Xiaoyan Yi¹, Zujing Zhang¹ & Weiping Yang¹

The rapid expansion of data centres has intensified the demand for low-threshold thermal optimisation strategies that avoid the high costs and operational disruptions associated with conventional infrastructure retrofits. This study focuses exclusively on in-rack server layout adjustments, specifically server power combination, arrangement methods, and air supply velocity. A dedicated experimental platform was constructed to represent a standard server rack environment, and physical single-factor and orthogonal experiments were conducted to evaluate cooling efficiency. Results indicate that: (1) Heat accumulates primarily at the rack top, producing pronounced temperature stratification. (2) The relative impact of the examined parameters on thermal performance ranks as follows: air supply velocity > power combination > arrangement method. (3) An optimal configuration—400 W × 10 servers positioned in the lower rack with an air supply velocity of 4 m/s—effectively alleviates heat imbalance and suppresses hot-spots. Compared with a non-optimised baseline, this configuration yielded measurable improvements in airflow utilisation and thermal uniformity. This work demonstrates that strategic in-rack layout adjustments alone can markedly enhance cooling performance without additional capital investment or infrastructure modifications, offering actionable guidance for data centre operators seeking practical, low-effort optimisation measures.

Keywords Fan wall cooling technology, Data centres, Thermal environment, Server arrangement, Thermal management

Abbreviations

CFD	Computational fluid dynamics
CRAC	Computer room air conditioner
PUE	Power Usage Effectiveness
IOM	Index of mixing
RTI	Return temperature index
SAV	Supply air volume
SHI	Supply heating index
T	Temperature
U	Unit

List of symbol

Re Reynolds number

Greek symbols

β	Coefficient of expansion [K^{-1}]
σ	Degree of dispersion of temperature
ρ	The density of air [kg/m^3]
S_{σ}^2	Sample variance
x_i	Measurement value obtained from the i th measurement, $i = 1, 2, 3 \dots n$
$\bar{x}$	Arithmetic mean of n measurements

¹College of Civil Engineering, Guizhou University, Guiyang 550025, China. ²School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield AL10 9AB, UK. ✉email: rymao@gzu.edu.cn; mec.chu@gzu.edu.cn

n Number of measurements
 σ Standard deviation

Driven by emerging technologies such as 5G and industrial IoT^{1,2}, data centres, as the core of digital infrastructure, have continued to grow in service scale and energy demand^{3,4}. According to the survey, the current energy consumption of data centres accounts for 4.2% of global energy consumption⁵, and this proportion is expected to reach 4.5% by 2025. Data centre temperature control systems, which are required year-round to maintain safe equipment temperatures due to high thermal loads, account for approximately 30–40% of total operational energy consumption^{6–8}. Studies have shown that cooling inefficiency and server overcooling directly lead to energy wastage in data centers^{9,10}. Consequently, thermal management optimisation is pivotal to data centre efficiency, driving research to reduce cooling energy use and safeguard the reliable operation of digital infrastructure.

The substantial energy consumption of the cooling system poses a key constraint to the sustainable development of data centers¹¹. To solve this problem, contemporary research primarily focuses on optimising the thermal environment at two levels: the server room and the rack. In data centres, changing the server room layout can improve airflow efficiency, enhance the thermal environment, and reduce energy consumption. One of the common approaches for thermal management of server rooms is to close the cold or hot aisles, which helps to prevent the mixing of hot and cold air and reduce problems such as hot air recirculation and cold air bypass^{12–14}. While both cold and hot aisle containment improve thermal management, hot aisle containment demonstrates more pronounced benefits in optimising the server room's thermal environment^{15,16}. The fan wall, which delivers cold air to the server room a pivotal role in enhancing the internal thermal environment. By constructing a similar model, Wu et al.¹⁷ found that changing the arrangement of the fan wall can improve airflow organisation, thereby reducing rack temperature. The configuration—comprising a fan array area of 4320 cm², a distance of 81 cm from the first rack row, and a height of 18 cm—demonstrated marked improvements in cooling metrics. Specifically, the supply heat index (SHI) showed a 64.10% increase, while the return temperature index (RTI) approached the optimal value of 1. Significant improvements were also observed in the index of mixing (IOM) and the β value, which increased by 64.44% and 57.50%, respectively. Zhang et al.¹⁸ constructed a similar model and found that, when the gap was 60% and the distance between the screen cylinder and the fan was set to 35 cm, the hot-spot temperature was effectively reduced, and rack temperature uniformity was improved. Wang et al.¹⁹ used computational fluid dynamics (CFD) simulations, validated against on-site experimental measurements, to examine the thermal implications of various airflow containment configurations in a financial data centre. It was revealed that adopting an overhead hot-aisle enclosure without side panels between adjacent racks substantially mitigates hot-air recirculation and eliminates localised hot spots behind the server cabinets. Omrani et al.²⁰ integrated a neural network-based surrogate model with a genetic algorithm to optimize inlet air velocity and temperature parameters in a fan-wall cooled medium-density data center. The proposed AI-driven optimization framework was demonstrated to achieve reductions in fan power consumption of up to 98.5% and in cooling load of up to 65.5%, thereby validating that dynamic adjustment of supply air conditions yields substantial improvements in energy efficiency while maintaining prescribed thermal safety margins.

Functioning as the critical thermal interface between the server cluster and the ambient environment, the rack's thermal environment reflects real-time operational characteristics of the equipment. It directly affects the heat distribution in the server room. Suboptimal airflow management results in non-uniform cooling airflow distribution, internal air recirculation patterns, and localised hot-spots, thereby elevating the operational load. Therefore, the spatial distribution of the rack, the number of devices, and the internal structure configuration are the key factors that determine the thermal environment of the rack^{21–23}. Wu et al.²⁴ proposed that installing baffles at the rack entrance can make the airflow uniform. They found that when the baffle's width is 7 cm, the angle is 15, the angle direction is the same as the airflow direction, and the rack will be more efficient when the baffles are installed in the same direction. Kuzay et al.²⁵ proposed a dynamic baffle deflector system installed at the rear of the rack, featuring an adjustable deflector that reduced equipment exhaust temperature by approximately 2.5 °C. Notable research was conducted by Lim's team²⁶, which developed a directional flow grill system that successfully suppressed temperature rise in hot-spot areas within safe thresholds using a 60° tilt configuration. Yuan et al.^{27,28} proposed tilting servers in a rack and showed that tilting servers at an angle of 30° angle helped dissipate between 5.7% and 10% more heat. This research team also^{28–30} constructed a tilted deflector structure at the rack entrance, equipment base, and floor mezzanine. Experimental results showed that this design improved airflow uniformity and reduced the extent of localised high-temperature regions. Jiao et al.³¹ employed a validated CFD model to evaluate the influence of server placement and to propose a deflector-based air supply strategy for in-row cooling architectures. Their findings indicated that distributing servers either at the top portion or uniformly throughout the rack, in conjunction with deflector angles ranging from 15° to 60° markedly enhances supply airflow uniformity within the cold aisle and lowers server intake temperatures. Although the above studies could optimize the thermal environment by adjusting the structure of racks and computer rooms, their practical implementation in operational environments presents significant engineering challenges, limiting widespread adoption.

While extensive research efforts have been directed toward thermal management in data centres, the preponderance of existing investigations has concentrated on room-level interventions—including but not limited to the optimisation of air supply technologies, the adjustment of computer room air conditioner (CRAC) setpoints, and the implementation of hot- or cold-aisle containment systems. Although such strategies have proven efficacious in newly constructed facilities or those with sufficient capital reserves, their deployment invariably requires infrastructure reconfiguration, non-trivial capital expenditures, and operational downtime. Consequently, these approaches remain largely inapplicable to the substantial inventory of legacy data centres that operate under stringent budgetary and continuity constraints.

This study innovatively proposes a fundamentally divergent paradigm: the strategic adjustment of in-rack server layout parameters without recourse to hardware acquisition or facility-level modification. Through controlled physical experimentation within a representative rack environment, the systematic manipulation of server power combination, arrangement methods, and air supply velocity was examined. It is demonstrated that significant amelioration of thermal maldistribution can be attained exclusively through intra-rack reconfiguration. This investigation thus addresses a conspicuous lacuna in the extant literature by providing actionable optimisation protocols specifically tailored to resource-constrained, ageing data centre infrastructure. Specifically, to systematically investigate the dynamic effects on internal rack airflow, single-factor experiments were conducted on three key parameters under rated-power conditions: server power combination, air supply velocity, and server arrangement methods. Subsequently, using an orthogonal experimental design, the thermodynamic effects of these parameters under multi-factor interactions were analysed to identify the optimal synergistic configuration scheme.

Methods

Experimental details

Experimental environment and principle

The model for this experiment was constructed based on the experimental setup described in Lan³². As shown in Fig. 1, the initial experimental environment consisted of a ventilated room (3.6 m × 3.1 m × 3 m) and a temperature-controlled room (2 m × 3.1 m × 3 m). A window connected the two rooms, measuring 1.1 m × 0.9 m. The ventilation room has four rectangular air vents located 2.2 m above the floor. These vents are connected to air ducts (0.5 m × 0.3 m) for air supply and exhaust to the outside, and a separate valve controls each duct. A refrigeration unit with a cooling capacity of 4.05 kW is installed in the temperature control room. The room's temperature range can be adjusted from −20 °C to 50 °C using a microcomputer controller connected to three PT1000 temperature sensors. Once the temperature range is configured, the air-conditioning unit automatically stops running when the indoor temperature exceeds the set limits and resumes operation when the temperature falls back within range.

Although simplified from Lan's original four-rack model³² to a single-rack setup due to practical constraints, this experimental design remains valid for studying thermal phenomena. Crucially, as evidenced by Wu¹⁷ and Song³³, hot-spot location is independent of rack quantity but is determined by the experimental configuration, exhibiting consistent patterns under identical conditions. This model thus provides a reliable platform for evaluating parameters affecting thermal performance. As shown in Fig. 2, a standard rack with dimensions of 0.6 m × 0.6 m × 2 m was placed in the room, and the hot-aisle outlet of the rack was sealed with 1 cm of insulation cotton.

The hot-air outlet of the rack was sealed with 1 cm of thick insulation cotton to form a closed exhaust channel measuring 0.6 m × 1 m × 2 m, thereby simulating the ventilation room as a data room. The server casing inside the rack was sized to a standard 2 Unit (U) server and made of galvanised steel with a thickness of 0.1 mm, a length of 44 cm, a width of 30 cm, and a height of 8.9 cm. Each server was equipped with heating rods of varying power to control its thermal environment. To better simulate uniform heat generation within the simulator, a snake-shaped heating rod measuring 43 cm in length and 15 cm in width was used in this experiment. To send

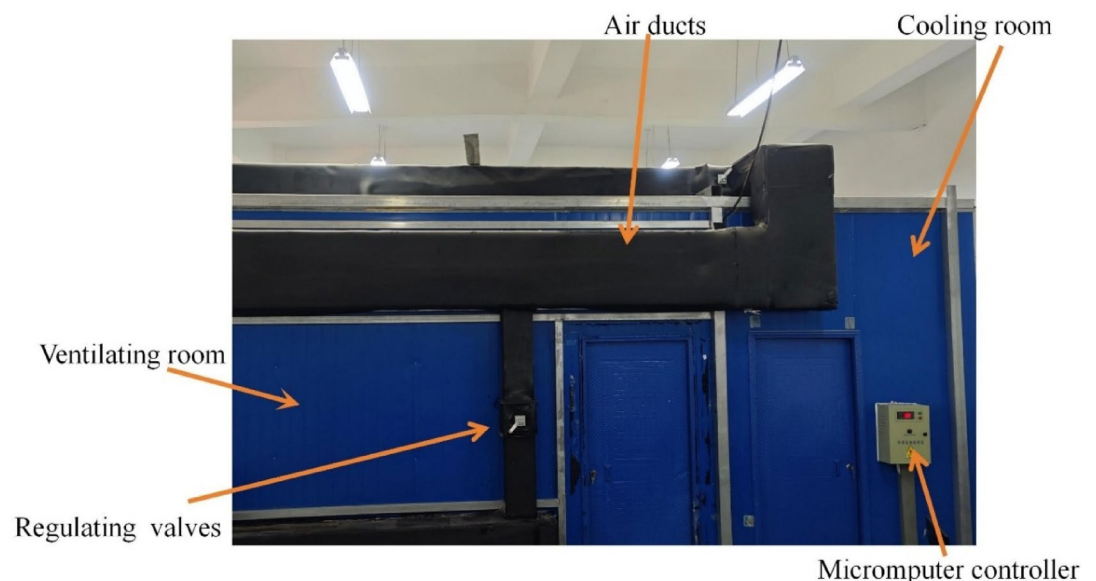


Fig. 1. Photograph of the experimental environment. The experimental setup consists of three main functional zones: the ventilating room, the cooling room, and the air ducts connecting them. Key components include regulating valves for airflow adjustment, and a microcomputer controller for parameter monitoring and system operation.

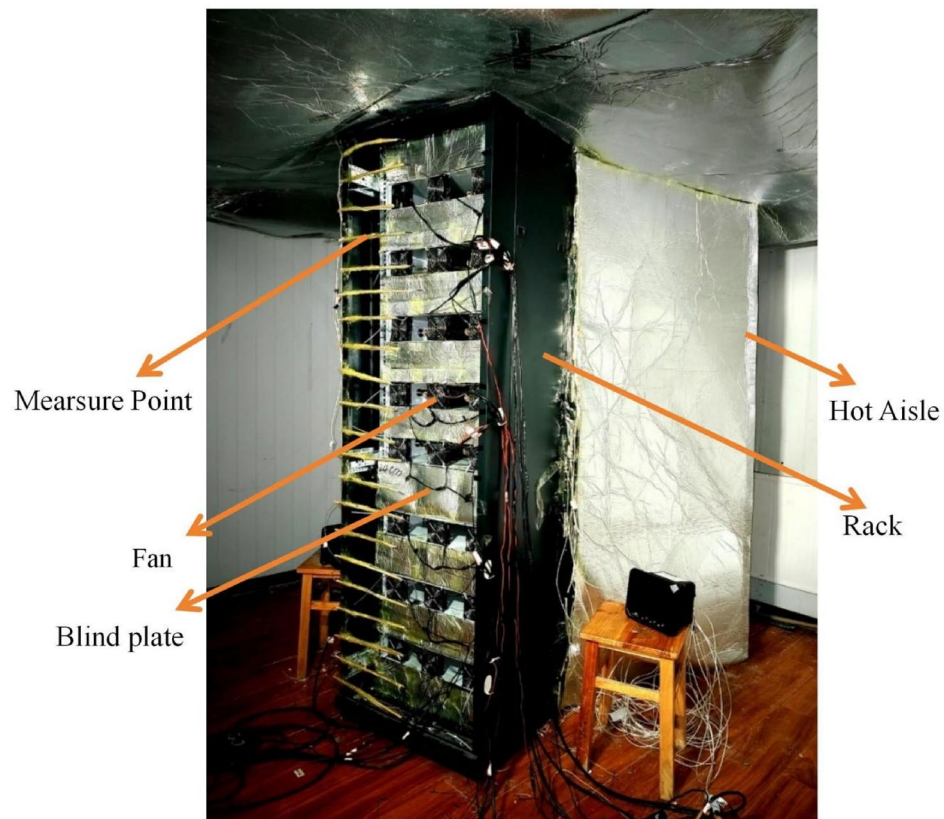


Fig. 2. Experimental platform of the cooling data center. The 42U server rack is equipped with simulated server units, fans, and blind plates. Temperature measurement points are arranged along the vertical direction of the rack, and the right side of the rack is connected to the enclosed hot aisle.

the heat generated by the server into the hot channel, three small square fans with an adjustable air velocity of 8 cm were installed at the air inlet of each server, controlled by the self-speed knob in the range of 0–5 m/s. The heat generated by the server was directed into the hot channel through the air inlet of each server. The gap between the two servers was set up as a blind plate with insulation foam to minimise the return of hot air and match the rack layout in the real data room. Four 0.6 m-diameter inverter fans were installed in the connecting windows of the data room and the heated rooms to simulate a fan wall system. The Supply air volume (SAV) of the fans was controlled between 0 and 3.5 m³/s using a frequency converter (model: 4.0G2-220 V). As shown in Fig. 3.

Figure 4 shows the laboratory's principal airflow process. Fresh outdoor air enters the cooling room and is cooled to the set temperature by the air conditioning unit. The cooled air then enters the data room, assisted by four variable-frequency fans. After the data room temperature stabilises to match the inverter air wall's exhaust air temperature, server fans push cold air into the servers. The air passes through the heating rods, becomes hot, and flows sequentially through the hot aisle and exhaust ducts before being discharged outside.

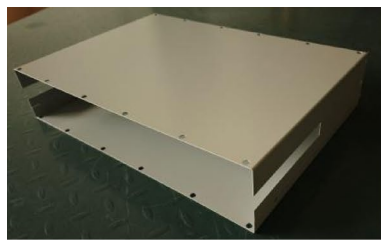
Arrangement of measuring points

The inverter fan wall air supply temperature and the inlet and outlet air temperatures of the server are essential parameters for calculating the data centre's thermal environment. For this reason, four temperature measurement points were set at the air outlet of the inverter fan, and the air supply temperature was obtained by calculating the average value. As shown in Figs. 5 and 6, 21 temperature measurement points were set up at the air inlet (cold aisle) and air outlet (hot aisle) of the rack, spaced 2U apart on one side, to determine the temperature distribution and accurately capture the inlet and outlet air temperatures of the servers. PT1000 thermocouples were used to measure air temperature at these points, with an accuracy of ± 0.01 °C. The MIK-6000F paperless data collector was used to record the temperatures. The air velocity of the inverter fan wall and the server inlet fans was measured using a FLUKE-925 handheld anemometer, with a measurement range of 0.40–25.00 m/s and an accuracy of 2%. To ensure the accuracy of the experimental data, the temperature and air velocity measurement equipment were debugged and calibrated before installing the experimental system.

Design of the experimental case

Definition of influencing factors

In actual data rooms, servers operate at different power levels. Since each rack has an upper power limit, variations in server power affect both the number of servers that can be installed and each server's heat dissipation.



(a) Server shells



(b) Heating rods



(c) Server cooling fan



(d) Variable frequency fan



(e) Frequency converters



(f) Wind speed meter

Fig. 3. Diagram of the central experimental apparatus. (a) Server shells for simulating server units. (b) Heating rods acting as heat sources to simulate server power consumption. (c) Server cooling fans. (d) Variable-frequency fans. (e) Frequency converters for airflow regulation. (f) a wind speed meter for measuring air velocity.

Therefore, it is essential to investigate how different power combinations influence overall heat dissipation. Moreover, it is commonly assumed in engineering practice that higher fan airspeeds lead to better cooling performance. However, it remains unclear whether higher airspeeds are always beneficial or whether they may instead induce vortices within the rack, reducing cooling efficiency. Furthermore, under full-load conditions, there is limited research on identifying the maximum airspeed that simultaneously satisfies cooling requirements and minimises energy consumption. For example, certain models, such as the Lenovo³⁴ ThinkSystem SR630 V2 (Types 7Z70, 7Z71), do not specify a maximum fan speed; instead, they offer only two fan options—standard and high-performance—selected based on the maximum ambient temperature, cooling conditions and the air baffle. Therefore, it is necessary to investigate the impact of the maximum wind speed at full load on the server's heat dissipation. Field investigations of actual data rooms have revealed that server layouts are often irregular, and different arrangement methods significantly affect the overall temperature distribution within the rack. In summary, this study focuses on the effects of three parameters on the cooling efficiency of a single data rack, namely: (A) the server power combination, (B) the air supply velocity, and (C) the server arrangement methods.

In real-world data centres, multiple server power configurations can coexist within a rack, while the rack's total power is typically fixed. Based on field research conducted in the data centres of the Guian New Area, this experiment selected four commonly used engineering server power ratings: 400 W, 500 W, 800 W, and 1000 W. Field surveys also showed that 4000 W is the most frequently observed total power for a single rack, which is consistent with the mainstream rack power configurations in this region. Accordingly, the experimental rack was assigned a rated total power of 4000 W. Four typical power combinations were tested: 400 W × 10, 500 W × 8, 800 W × 5, and 1000 W × 4. All combinations maintain a total power of 4000 W and cover both low-power high-density and high-power low-density layouts.

Since server fan airspeeds vary significantly across brands and models, making direct comparisons difficult, this experiment adopted four representative airspeeds: 2 m/s, 3 m/s, 4 m/s, and 5 m/s.

Field investigations have shown that server layouts in data centres are often disorganised, with uniform arrangements being the most common. However, few studies have systematically evaluated which configurations are most conducive to heat dissipation. To more realistically investigate the impact of server arrangement patterns, this study designed and tested the four most common server arrangement modes in actual data centres: upper, middle, lower, and uniform distribution. Specifically, these four arrangement schemes are shown in Table 1.

In summary, this study investigated the effects of three key factors on rack cooling efficiency and analysed their interactions under a constraint on constant total rack power. The findings provide practical guidance for optimising server power combinations, improving exhaust airflow strategies, and selecting appropriate server

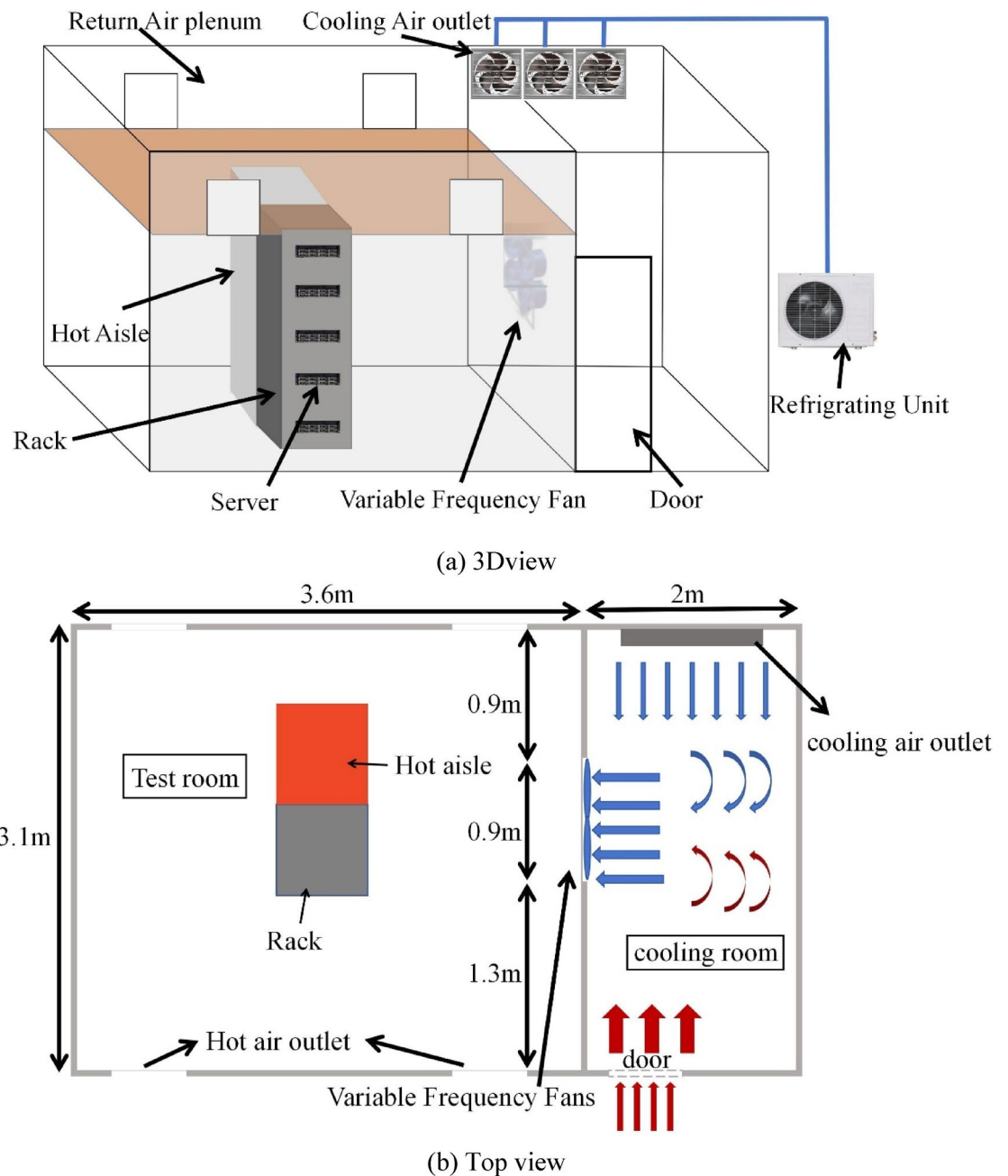


Fig. 4. DC Layout. a, (a) 3D view illustrating the overall structure, including the test room with rack and hot aisle, cooling room with variable-frequency fans and cooling air outlets, and external refrigerating unit. (b) a top view showing the dimensions of the test room (3.6 m × 3.1 m) and cooling room (2.0 m × 3.1 m), along with the arrangement of the rack, hot aisle, variable-frequency fans, and airflow paths.

arrangement methods in real-world data centre operations. These strategies can enhance thermal performance at minimal cost.

Initial working conditions

After field research, it was found that 400 W servers have the highest proportion in practical engineering applications. Uniform distribution was identified as the most widely used method among the possible server arrangements. Due to variations in fan air speed across server models and brands, the initial experimental wind speed was set to 3 m/s. Accordingly, the initial working condition was defined as follows: ten 400 W servers uniformly distributed within the rack, with an inlet air speed of 3 m/s.

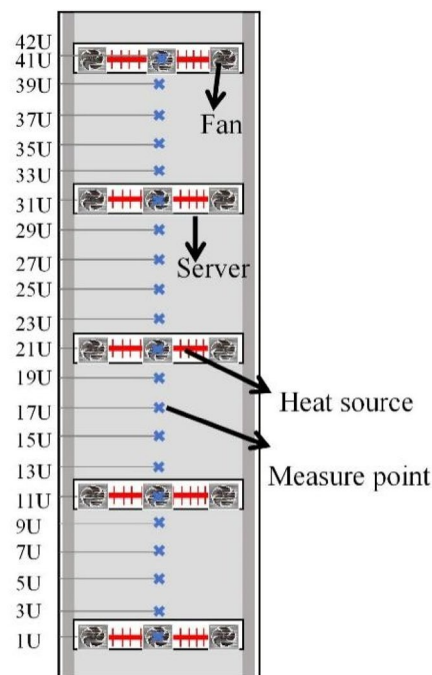


Fig. 5. Distribution of measurement points. Front view of the 42U server rack, showing the vertical arrangement of server units, fans, heat sources, and temperature measurement points (blue crosses) at different U positions from 1 to 42U.

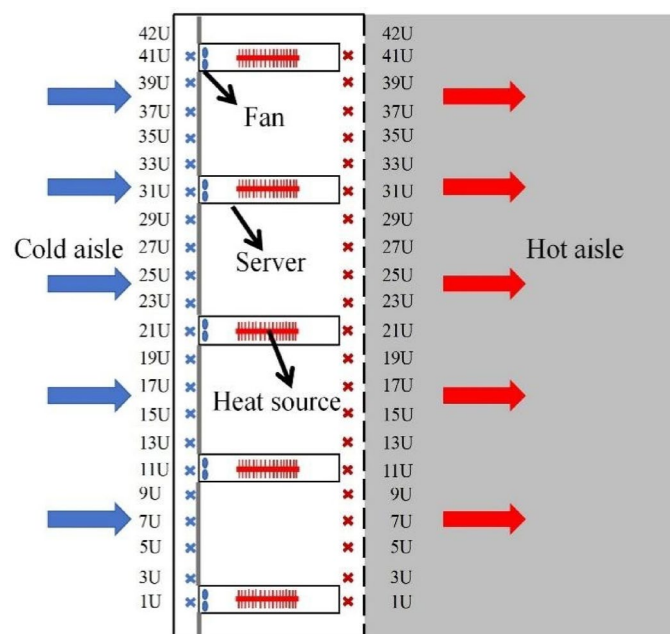


Fig. 6. Side view of measurement point. The side view of the 42U rack, with arrows indicating the path of the airflow and crosses representing the temperature measurement points. Blue represents the cold channel side, and red represents the hot channel side. Servers are distributed within the rack, and there are heat sources (red) inside. The server is equipped with small fans (blue) in the cold aisle.

Single-factor experiment

A single-factor experiment was conducted to evaluate the influence of individual variables on the rack's heat dissipation performance. Three factors were investigated, each with four levels determined according to typical engineering conditions. Each factor was subsequently analysed to determine the optimal parameter settings based on the experimental results, as shown in Table 2.

Server power combination	Upper arrangement	Middle arrangement	Lower arrangement	Uniform arrangement
400 W × 10	Server:14–42 Vacant:1–13	Vacant:37–42 Server:8–36 Vacant:1–7	Vacant:20–42 Server:1–19	Vacant:40–42 Server:2–39 Vacant:1
500 W × 8	Server:20–42 Vacant:1–19	Vacant:34–42 Server:11–33 Vacant:1–10	Vacant:24–42 Server:1–23	Vacant:40–42 Server:3–39 Vacant:1–2
800 W × 5	Server:39–42 Vacant:1–38	Vacant:29–42 Server:15–28 Vacant:1–14	Vacant:15–42 Server:1–14	Server:42–1(full)
1000 W × 4	Server:32–42 Vacant:1–31	Vacant:29–42 Server:18–28 Vacant:1–17	Vacant:12–42 Server:1–11	Server:42–1(full)

Table 1. Specific locations of servers in different arrangements. The rack's total height is 42U. A 2U spacing is maintained between servers in non-uniform arrangements. For uniform arrangements, servers are installed consecutively.

Experimental case	Parameter	A(W)	B(m/s)	C
1	Server power combination	400 × 10	3	Uniform arrangement
2		500 × 8		
3		800 × 5		
4		1000 × 4		
5	Air supply velocity	400 × 10	2	Uniform arrangement
6			3	
7			4	
8			5	
9	Server arrangement methods	400 × 10	3	Lower arrangement
10				Upper arrangement
11				Middle arrangement
12				Uniform arrangement

Table 2. One-factor experimental design.

Orthogonal experiment

Orthogonal experiments were conducted to investigate the effects of various factors on the rack's thermal performance to identify the optimal working conditions. The three primary factors analysed were (A) the server power combination, (B) the air supply velocity, and (C) the server arrangement methods, each with four levels. A three-factor table with three columns was selected from the L16(4³) orthogonal experiment table for the three-factor, four-level experiment, as shown in Table 3.

Performance indicators

During the air supply to the server, if the wind speed is too high, it can easily trigger heat reflux. At the same time, the server's diverse layout and power combinations have a non-negligible impact on the airflow organisation of the hot channel and, in severe cases, even cause airflow short circuits, which not only significantly reduces the thermal efficiency of the server but also leads to local hot spots in the rack, threatening the stability of the server's operating status. In the previous experimental conditions, the thermal channel of the rack was not fully closed, and a return flow of hot air in the upper part of the rack affected the temperature at the corresponding air inlet, leading to hot spots. The current SHI³⁵, IOM³⁶, β -value³⁷, and a series of evaluation indices were based on this particular working condition. However, in this experiment, the cold-channel inlet of the rack created positive pressure via the fan to prevent the recirculation of hot air. A blind plate was placed between adjacent servers to isolate the hot channel. Closing the hot channel prevented the recirculation of hot air from above the rack into the cold aisle, effectively isolating the hot and cold air streams. Therefore, evaluating and analyzing according to this kind of index caused a contradiction between the actual hot aisle temperature distribution and indices.

To scientifically evaluate thermal performance and mitigate server damage caused by uneven airflow and hot-spots, this study, following Wang³⁸, introduced the degree of discreteness, an index of the uniformity of the hot aisle temperature distribution, to comprehensively assess thermal efficiency. The degree of discreteness quantifies deviations from the average temperature, reflecting the uniformity of the server's temperature distribution. This metric serves as an indicator of the balance between internal heat dissipation. A lower degree of discreteness means the server temperature distribution is more uniform, and the cooling system operates more stably. In contrast, a large σ indicates severe non-uniformity in server outlet airflow, causing localised heat accumulation that cannot be removed in time and consequent abnormal temperature rises at corresponding outlet positions. This non-uniformity and abnormal temperature elevation create two major hazards for servers. First, reduced

Experimental case	A(W)	B(m/s)	C
13	400×10	2	Lower arrangement
14	400×10	3	Upper arrangement
15	400×10	4	Middle arrangement
16	400×10	5	Uniform arrangement
17	500×8	3	Lower arrangement
18	500×8	2	Upper arrangement
19	500×8	5	Middle arrangement
20	500×8	4	Uniform arrangement
21	800×5	4	Lower arrangement
22	800×5	5	Upper arrangement
23	800×5	2	Middle arrangement
24	800×5	3	Uniform arrangement
25	1000×4	5	Lower arrangement
26	1000×4	4	Upper arrangement
27	1000×4	3	Middle arrangement
28	1000×4	2	Uniform arrangement

Table 3. Table of orthogonal experiments L16(4³).

service life: hot air accumulating on the exhaust side continuously transfers heat back to internal components, and excessively high outlet temperatures imply that heat removal is insufficient to offset generation. Following the Arrhenius model³⁹, a roughly 10 °C increase in component operating temperature can halve expected lifetime, with sudden rises exponentially accelerating aging and failure. Second, increased airflow resistance and a vicious cycle of reduced airflow: higher outlet temperatures lower air density and raise dynamic viscosity, demanding a larger driving pressure head to maintain flow while increasing frictional losses. This forces server fans away from efficient operating points, reducing cold air delivery, which further elevates outlet temperature—a cycle of “temperature rise → airflow reduction → even higher temperature rise.” In severe cases, thermal runaway can sharply degrade computational performance.

To better evaluate the server’s thermal environment, this study analyses optimal parameter values for each factor using a two-step process. First, it determines whether the temperature at the server outlet is excessively high. Subsequently, it checks whether σ meets the specified requirements. Both steps are indispensable. In short, the lower the degree of discreteness, the smaller the temperature difference between server outlets, the more uniform the temperature distribution, and the better the server’s thermal performance.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}} \tag{1}$$

where, x_i is the temperature of each server outlet, $\bar{x}$ is the average server outlet temperature, and n is the number of servers.

Experimental process

In this experiment, 21 temperature measurement points were placed at the hot and cold channel positions on the rack. The MIK-6000F data acquisition instrument was used, and records were collected every 10 s. The air supply speed was adjusted real time via the frequency converter (4.0 G2-220 V) to maintain a rate of 2.8±0.2 m/s, and the air supply temperature was maintained at 23 ± 1 °C. The air supply speed was referred to in the experimental model of Lan³², but the air supply speed of the wind wall was also increased because the total power of the rack was increased in this experiment. The air supply temperature was taken according to the average summer temperature of 23 °C in Guiyang. Different working conditions were achieved by varying the server’s power combination, arrangement methods, and air supply velocity during the experiment. The specific experimental process is as follows:

1. Activate the cooling room’s control system. Set the cooling room temperature to 23 °C and the variable-frequency fan speed to 2.8 m/s.
2. Commence the experiment once the cold aisle temperature stabilises to match the supply air temperature from the wind wall.
3. Ensure that the server power combinations, arrangement methods, and air supply velocity are correct, then power on the servers and server fans.
4. Configure the data logger to record temperature data at 10-s intervals. Continue data acquisition until both hot and cold aisle temperatures remain stable for approximately 30 min.
5. Observe the data collected by the data logger. Once data logging is complete, turn off the servers.
6. Once the data room temperature returns to its initial conditions, adjust the operational settings and begin a new group of experiments.
7. Upon completion of all experimental trials, extract the data from the logger for calculation and analysis.

At the end of each group of experiments, cut off the heating bar's power supply while keeping the fan running continuously. This can effectively improve the server's cooling rate and shorten the system's cooling time. All the experimental steps are presented in the Fig. 7.

Uncertainty analysis

The uncertainty of experimental data can be calculated using the method proposed by Moffat⁴⁰. The measurement uncertainty is calculated using the following formula:

$$u = \sqrt{(u_A)^2 + (u_B)^2} \quad (2)$$

where u is the uncertainty of the direct measurement result, u_A is the Type A uncertainty caused by measurement repeatability, and u_B is the Type B uncertainty caused by the error of the measurement device.

For conventional measurements under repeatability or reproducibility conditions, it is usually unnecessary to evaluate the Type A standard uncertainty for each measurement, and the pre-evaluated results can be directly adopted. The standard deviation of a single measurement of the measured parameter is calculated using the Bessel formula (Eqs. (3) and (4)).

$$S_\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (3)$$

$$u_A = \frac{S_\sigma}{\sqrt{n}} \quad (4)$$

After conducting 10 independent repeated measurements in advance, the standard deviation of a single measurement is calculated using the above formula, yielding $u_A = 0.28^\circ\text{C}$.

In this experiment, a Class A precision PT1000 temperature sensor with a measurement range of -50°C to 200°C and an allowable reading error of $\pm 0.01^\circ\text{C}$ is used. The measurement error follows a uniform distribution; therefore, the confidence factor $k = \sqrt{3}$ and the half-width of the interval $a = 0.01^\circ\text{C}$ are adopted. Accordingly, the Type B standard uncertainty u_B introduced by the temperature measurement error is obtained as:

$$u_B = \frac{a}{k} = \frac{0.01^\circ\text{C}}{\sqrt{3}} \approx 0.0058^\circ\text{C} \quad (5)$$

Substituting $u_A = 0.28^\circ\text{C}$ and $u_B = 0.006^\circ\text{C}$ into Eq. (1), the uncertainty of the measured temperature is obtained as $u = 0.28^\circ\text{C}$. The measurement accuracy is relatively high, ensuring the reliability of the experimental results.

Results and discussion

Temperature distribution of the initial experimental group

The initial working condition is set to 10 servers, with a single power of 400 W, evenly distributed, spacing 2U, and a supply air velocity of 3 m/s. As shown in Fig. 8, the rack outlet temperature shows a significant vertical gradient, with an overall temperature difference of 20.42°C . Specifically, the temperature at the exit of each server increases with height, with the 9th server exhibiting an abrupt increase and an inflexion point, and the growth rate slows at the top of the 10th server.

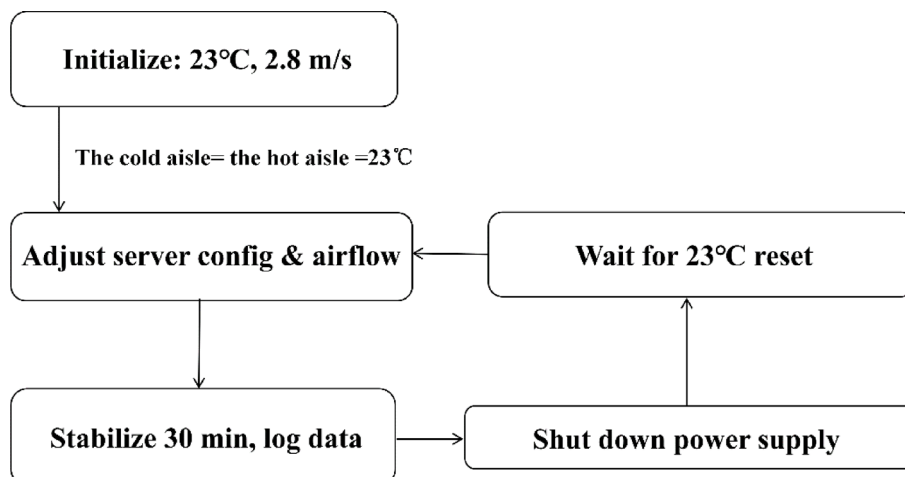


Fig. 7. Experimental procedure.

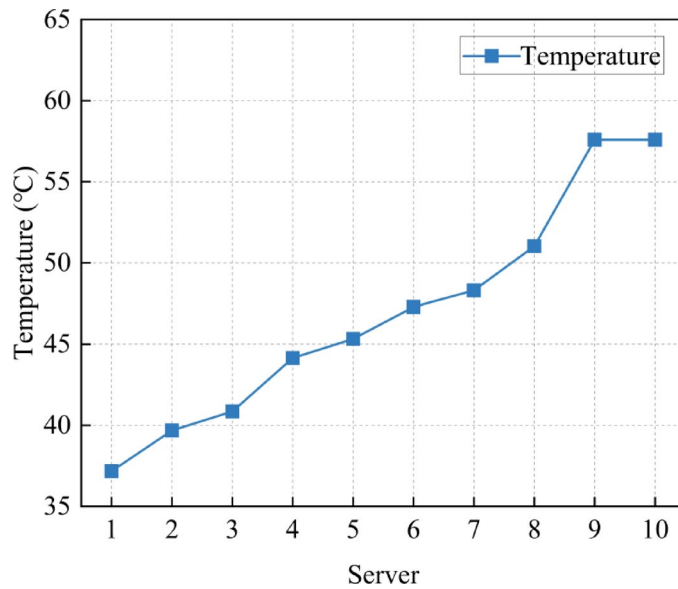


Fig. 8. Initial operating temperature graph.

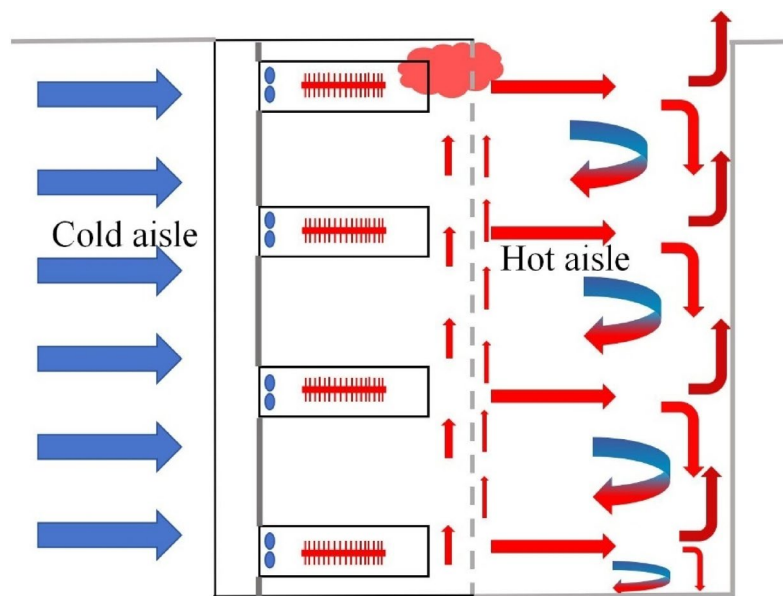


Fig. 9. Air flow organisation movement diagram. Side view of the rack showing the airflow path. The arrows indicate the direction of air flow, and the colors represent the temperature.

As shown in the Fig. 9, when servers exhaust hot air at high velocity, the jets impinge on the rear hot-aisle containment wall, converting dynamic pressure to static pressure. This creates a localised high-pressure zone that bifurcates the airflow: one branch ascends along the wall, while the other is deflected downward into the lower rack volume. According to the ideal gas law ($\rho \propto 1/T$), the downward-directed branch is less dense than the surrounding cooler air and therefore experiences an upward thermal buoyancy force. Since the rack's fan or air-conditioning airflow is typically insufficient to overcome this buoyancy, the hot air cannot be sustained in the lower region; instead, it rises back, merges with the exhaust from servers at higher elevations, and progressively elevates their temperatures. At the same time, the narrow physical space at the top of the rack restricts the diffusion of the accumulated hot air, preventing it from being effectively discharged through natural convection or forced airflow. Consequently, heat builds up at the rack top, forming a pronounced hot spot. Given identical power dissipation per server, the exhaust air temperature of each unit would theoretically remain approximately uniform. Experimental observations, however, reveal a completely different trend: servers placed at higher elevations exhibit substantially elevated exhaust temperatures. If this phenomenon were attributable solely to direct convective heat dissipation from individual servers, such a temperature stratification

	Case7	Case15	Case19	Case23	Case27
Bottom server temperature (°C)	38.77	42.025	43.61	29.53	49.32
Lower measurement Point temperature (°C)	33.45	41.14	43.20	27.7	30.00
	31.28	37.82	42.33	26.49	26.56
	29.69	38.99	40.96	25.8	26.49
			40.67	25.52	25.74
			40.44	24.98	25.30
			40.87	24.47	24.95
				24.21	24.61
					24.49

Table 4. Temperature changes in the middle arrangement.

	Case6	Case18	Case22	Case26
Bottom server temperature (°C)	40.21	33.70	34.13	47.14
Lower measurement point temperature (°C)	37.30	27.96	30.47	42.96
	35.64	26.99	29.28	33.77
	31.87	25.81	26.95	30.11
	26.12	25.92	26.34	26.25
	25.44	24.45	25.52	24.60
		25.16	25.01	24.07
		25.02	24.73	24.45
		24.65	24.40	24.12
		24.53	24.17	23.65
				23.62
				23.36
				23.12
				22.94
				22.65
				22.27

Table 5. Temperature variation of the upper arrangement.

would violate the principle of energy conservation. This discrepancy indicates that heat is not fully exhausted but instead progressively accumulates within the rack enclosure. The presence of thermal recirculation is further corroborated by temperature measurements taken beneath the operating servers, as documented in Tables 4 and 5: all monitoring points below the lowest server registered temperature increases of 0.41 °C to 19.32 °C relative to ambient conditions, confirming substantial penetration of hot recirculated air into the lower volume. Therefore, the hot spot phenomenon results from the combined effects of thermal recirculation, thermal buoyancy, and the rack's geometric limitations.

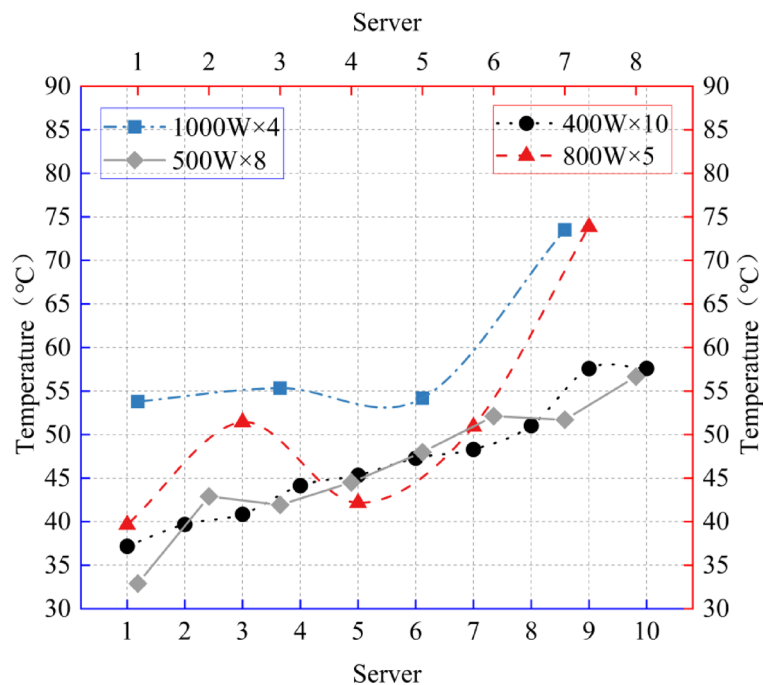
It is worth noting that in the experiments of Li⁴¹, the hot spot exhibited by the rack is in the lower part, which is because the blind isolation of the rack air inlet and the fan at the server air inlet in a real data room were not considered in the experimental model of the Li study. When the blind and fan are not installed, the open structure allows cold airflow to flow without practical guidance, and part of it is directly mixed with the hot air in the upper part of the rack, forming an unorganised circulating airflow. At the same time, the low-pressure area generated when the lower server inhales the cold air triggers the chimney effect, reversing the flow and drawing the top hot air back downward, so the lower server inhales the mixed airflow. After installing the blind and fan, the physical isolation of the blind and the positive pressure created by the fan allow the cold air to penetrate the server horizontally, eliminating the mixing of hot and cold airflow. The experimental model at Lan³², with the fan installed at the server's air intake, significantly reduces heat return, so the server outlet temperature trend is consistent with the experiment, verifying its reliability.

One-factor experimental analysis

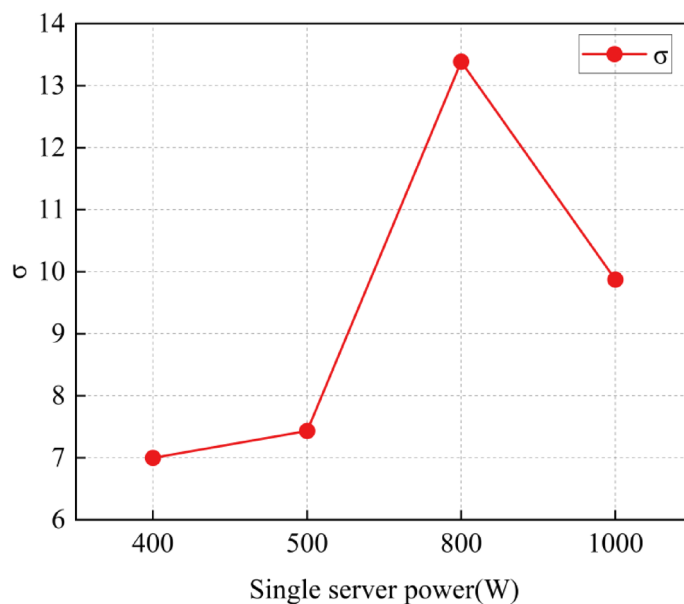
Server power combination

The experimental conditions were configured as follows: the air supply temperature was maintained at 23 °C, and both the room and server air supply speeds were set to 3 m/s. The power levels of individual servers were regulated by adjusting the heating rods to 400 W, 500 W, 800 W, and 1000 W, respectively. Figure 10a and b illustrate the outlet temperature distribution and corresponding discreteness under different power configurations.

If no other factors are considered, the ideal outlet temperatures for the 400 W × 10, 500 W × 8, 800 W × 5, and 1000 W × 4 configurations would be 28.8 °C, 30.2 °C, 34.5 °C, and 37.4 °C, respectively; however, the actual



(a) Server outlet temperature



(b) Temperature discreteness

Fig. 10. Different power combinations.

situation was completely different. As shown in Fig. 10a, under the condition of the rated power of 4000 W, the influence of different server power combinations on the server outlet temperature presents the following pattern: under the same total power of 4000 W configuration, the server temperature change presented a significant grouping feature: the 1000 W×4 and 800 W×5 combinations and the 500 W×8 and 400 W×10 combinations showed similar temperature trends, respectively. This was due to the slight power differences within each group, a similar number of devices, and similar server spacing and experimental environment conditions when evenly distributed. On the contrary, between the 1000 W/800 W high-power group and the 500 W/400 W low-power group, there was a significant difference in the power difference between the single machines and the density of the servers, which ultimately led to a considerable divergence in the temperature change law.

The overall temperature difference of the rack with the increase in stand-alone power showed a trend of first rise and then fall; when the stand-alone power increased to 800 W, it reached a peak of 34.19 °C; and when the stand-alone power increased to 1000 W, the temperature difference was significantly reduced to 19.72 °C. Because the thermal cycle causes the server to be placed at a higher position, resulting in more accumulated heat. Increasing the server power directly increased the heat flux, thereby increasing the spatial scale required for the downstream hot airflow to adequately mix with the ambient cold air. When the server power was ≤ 800 W, the space beneath the servers was insufficient to allow the hot exhaust airflow to dissipate heat through adequate mixing with the surrounding cold air. Meanwhile, this airflow was buoyantly lifted by the underlying hot air, progressively intensifying the heat buildup. Furthermore, the server power was positively correlated with the accumulation of undissipated heat, leading to an increase in the vertical temperature difference across the rack and its continued expansion. Consequently, the temperature difference reached its maximum at a single-server power of 800 W. When the power increased to 1000 W, the sufficient airflow passage space enabled the downstream hot airflow to fully mix, effectively suppressing heat accumulation; thus, the overall temperature difference was significantly reduced to 19.72 °C. Overall, the rack temperature was lower at low server powers than at high powers, and the temperature rise was smoothest at 400 W. Overall, different power combinations affect rack heat dissipation primarily through two factors: the heat output of individual servers and the airflow uniformity associated with the number of servers.

As shown in Fig. 10b, the degree of discreteness first increases with server power and then decreases at 800 W. A greater degree of discreteness indicates more uneven heat distribution. When the server power was ≤ 800 W, the increase in power intensified heat buildup at the top of the rack, leading to a larger overall temperature difference and, consequently, a higher degree of discreteness, which is consistent with the analysis in the previous section. Although the overall rack temperature difference was smallest at 1000 W, the high power resulted in higher heat dissipation intensity from individual servers, leading to excessively high top temperatures and increasing the degree of discreteness. Such high temperatures can adversely affect normal server operation. Therefore, considering the overall rack temperature difference, the upper operating temperature limit of the servers, and the degree of discreteness, the 400 W $\times$ 10 configuration was identified as the optimal combination among the tested single-server power options.

Air supply velocity

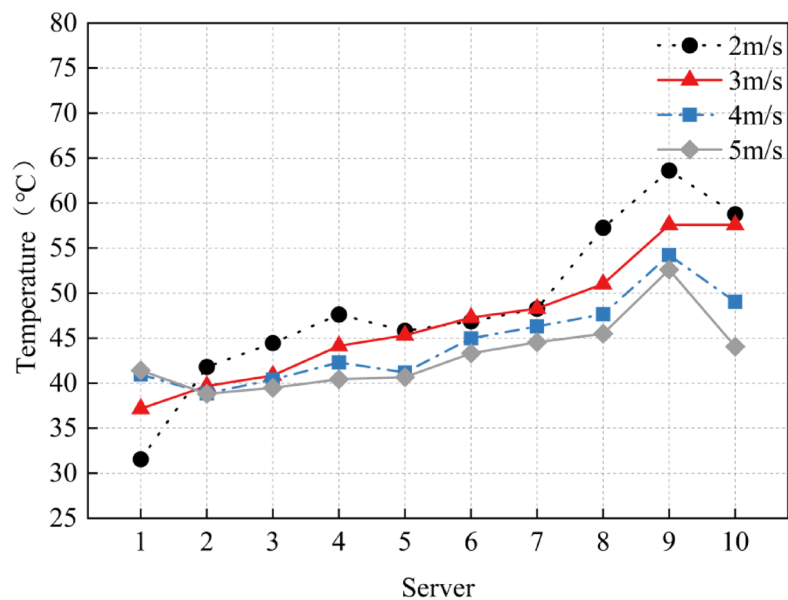
The experimental conditions were set as follows: the air supply temperature was set to 23 °C, the air supply velocity was 3 m/s, the power of a single server was 400 W, and the air supply velocity of the server was adjusted by adjusting the small fan velocity of the server air inlet to 2 m/s, 3 m/s, 4 m/s, and 5 m/s. Figures 11a and b show the discreteness of server outlet temperature vs temperature at different air supply velocities, respectively.

As shown in Fig. 11a, the effect of varying the server's air supply velocity on its outlet temperature is shown for different vertical heights of measurement points. In terms of temperature change planning, regardless of air supply velocity, the overall trend followed an upward gradient, and all showed the highest temperature at the top, consistent with the conclusion of 4.1. In addition, as air supply velocity increased, the rack's overall temperature gradually decreased. When the server air supply velocity reached 5 m/s, the whole rack reached a minimum in the hot-spot temperature and the overall temperature difference. This occurred because the high-speed flow attenuated the thermal boundary layer at the heating rod surface and increased the convective heat transfer coefficient, thereby improving cooling efficiency. Meanwhile, turbulence induced by high-velocity airflow enhanced mixing of hot and cold air, thereby reducing the local temperature difference. At the same time, high-speed airflow can overcome the traditional low-speed difficulty of reaching the "dead zone," such as the upper layer of the rack, to inhibit the formation of hot spots. In short, the increase in air supply velocity through enhanced forced convection and optimised airflow distribution were the two aspects of the coupling effect that achieved overall cooling and reduced temperature differences.

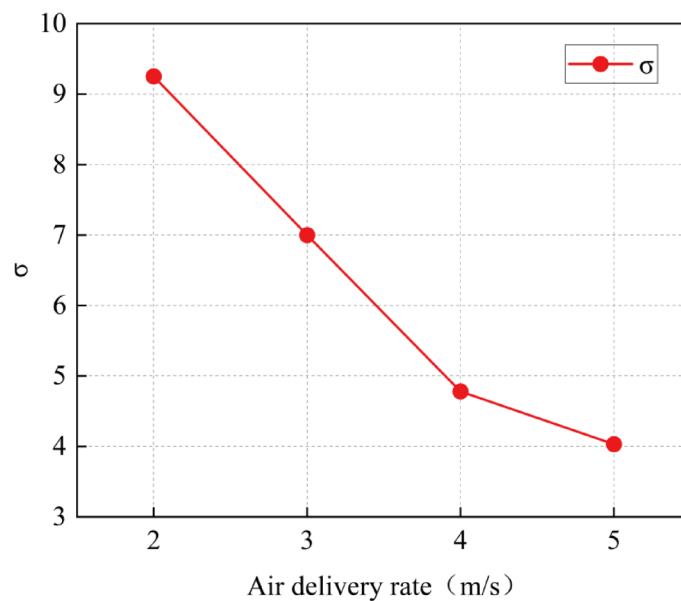
As shown in Fig. 11b, the degree of discreteness is negatively correlated with the air supply velocity, and the degree of discreteness gradually decreases with the increase of the server air supply velocity. It could be assumed that, with a low-speed air supply, the airflow Reynolds number decreased, leading the flow to transition to a laminar state. Concurrently, the cold air's inertia decreased, hindering its ability to penetrate deeply into the rack. This resulted in insufficient airflow distribution toward the far end of the dead zone, leading to the formation of a local high-temperature region. Ultimately, these factors contributed to an expansion of the temperature difference. The low-speed air supply's convection heat transfer was weakened, and heat accumulation on the equipment's surface affected the surrounding air temperature, resulting in deviations from the average value at each measurement point. Therefore, the airflow organisation and convective heat transfer, under the joint effect of the two aspects, led to the degree of discreteness increasing with decreasing wind velocity. Among them, when the wind velocity was 5 m/s, the degree of discreteness was 0.9580, which was 56.41%, 42.38%, and 15.65% lower than the other three wind velocities. It was observed that the increase in wind velocity and the reduction in temperature discreteness exhibited a nonlinear relationship. The discreteness at 4 m/s and the average outlet temperature showed little difference compared to those at 5 m/s, with only a 3% discrepancy between the two conditions.

To better analyse the energy-efficiency benefits of different wind speeds from an energy-consumption perspective, this paper introduces Power Usage Effectiveness (PUE) as an energy metric⁴², with its calculation formula shown in Eq. (6). As shown in Table 6, when the supply air velocity is 4 m/s, PUE = 1.0055; when the supply air velocity is 5 m/s, PUE = 1.0094. Therefore, with careful consideration of energy consumption, 4 m/s was the most economical air supply velocity. If only to reduce the overall temperature of the rack, 5 m/s was the best air speed.

Therefore, with careful consideration of energy consumption, 4 m/s was the most economical air supply velocity. If only to reduce the overall temperature of the rack, 5 m/s was the best air speed.



(a) Server outlet temperature



(b) Temperature discreteness

Fig. 11. Different server supply air velocity.

Case	Air supply velocity (m/s)	Total air volume (m3/s)	Import and export average temperature difference (°C)	Heat dissipation (kW)	Total power of the fan (W)	Rack PUE
Case 5	2 m/s	0.30144	21.592	7.849483131	9.3	1.0012
Case 6	3 m/s	0.45216	22.3735	12.20038052	31.2	1.0026
Case 7	4 m/s	0.60288	18.5225	13.46721483	73.8	1.0055
Case 8	5 m/s	0.7536	16.892	15.35215231	144	1.0094

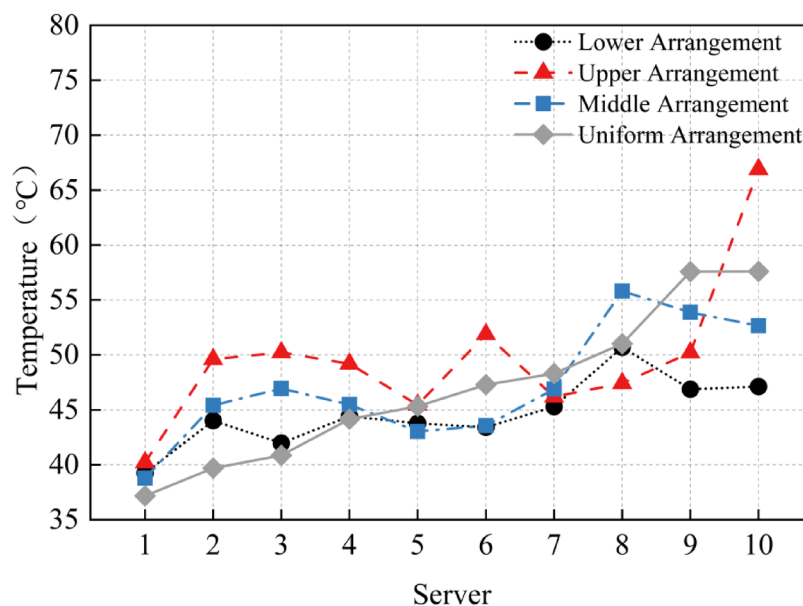
Table 6. PUE under different air supply velocities.

$$PUE = \frac{P_{Total}}{P_{IT}} \quad (6)$$

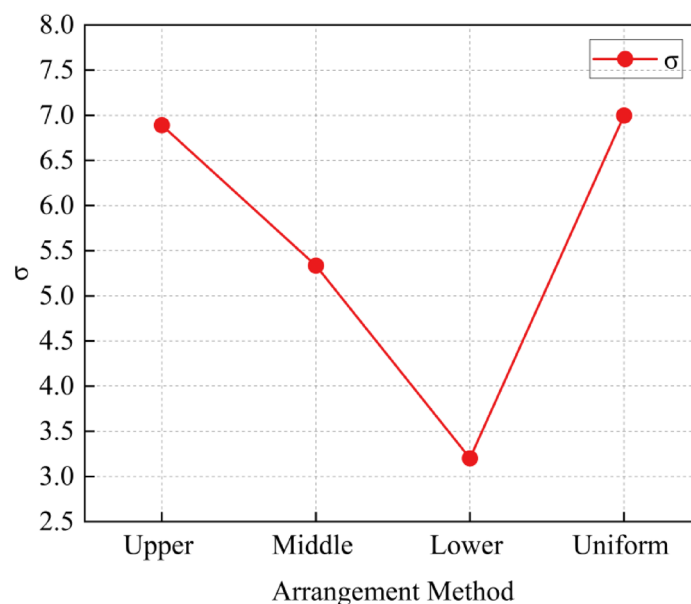
where P_{Total} is the total power consumption of the experimental setup (W), measured at the input and including the IT load, the cooling fans. P_{IT} is the power consumed exclusively by the IT equipment (W), held constant at all tested Air supply velocities.

Server arrangement methods

The experimental conditions were set as follows: the air supply temperature was set to 23 °C, the air supply speed was 3 m/s, the power of a single server was 400 W, the server cooling air speed was 3 m/s, and the server arrangement was adjusted by the spacing between the servers, as shown in Table 1. Figures 12a and b showed the discreteness of the server exit temperatures vs temperatures for the different server arrangement methods, respectively.



(a) Server outlet temperature



(b) Temperature discreteness

Fig. 12. Different arrangement methods.

As shown in Fig. 12a, different arrangement methods affect the server outlet temperature differently. The highest server temperatures consistently occurred in the upper-middle region, consistent with the conclusion in Section “Temperature distribution of the initial experimental group”. There was little difference in the bottom server outlet temperatures among the arrangements; the primary factor determining the overall temperature difference was the maximum temperature at the top. Both the maximum server temperature and the overall rack temperature difference followed the same trend: upper arrangement > uniform arrangement > middle arrangement > lower arrangement.

The high server density in the uniform arrangement increased the amount of insufficiently mixed hot air, leading to severe heat accumulation at the top and elevated hot-spot temperatures. The server density in the upper arrangement was even higher, resulting in a larger rack temperature difference than in the uniform arrangement. In both the middle and lower arrangements, the topmost servers had a vacant area above them, allowing the rising hot airflow to spread and dissipate rather than accumulate at the top outlets. The thermal data in Table 7 support this explanation: for the middle arrangements, the average temperature of the upper vacant area exceeded that of the lower vacant area, with the temperature differences across test cases ranging from 0.5 °C to 29.37 °C. For the lower arrangement, the temperature difference between the upper and lower vacant areas was smaller, ranging from −0.5 °C to 15.30 °C. The upper vacant zone in the lower arrangement was larger than that in the middle arrangement, providing more space to absorb diffused heat. Consequently, the overall temperature in the lower arrangement was lower and more uniform, and the rack temperature difference was the smallest among all configurations.

Figure 12b shows the effect of different server arrangements on the degree of temperature discreteness. The figure showed that the lower arrangement’s temperature discreteness was drastically reduced compared to the other groups, indicating that it led to a more uniform temperature at the exit of all servers and a significant reduction in heat buildup. This conclusion corroborated the temperature graphs presented by the different server arrangement methods. Therefore, considering the temperature uniformity and hot-spot temperature, the best server arrangement method was the lower arrangement.

Analysis of orthogonal experiments

In the previous paper, the power combination of 400 W × 10, a wind speed of 3 m/s, and a uniform layout were used as the initial conditions to investigate the effects of single-factor changes on the cooling performance of a single rack. However, it had not been possible to examine the comprehensive impact on the rack’s thermal environment from the simultaneous changes in three factors: (A) power combination, (B) air supply velocity, and (C) arrangement method. Therefore, orthogonal experiments were designed to verify the interaction effects among factors, explore the primary and secondary relationships among factors with respect to the evaluation indexes, and determine the optimal operating conditions. The degree of discreteness was used to evaluate the rack’s airflow organisation and thermal performance. The closer the degree of discreteness was to 0, the lower the volatility of the experimental data, the more uniform the airflow organisation, and the lower the likelihood of hot spots in the server under working conditions. Based on the L16(4³) orthogonal table, this study systematically explored the influence of three four—level factors, namely, server power combination (A), air supply velocity (B), and server arrangement methods (C), on the thermal performance of racks. The rack temperature data were collected through 16 groups of complete factorial experiments. The key index of discreteness was calculated, as shown in Table 8. Then, the factors were ranked in order of significance using range analysis, as shown in Table 9. Subsequently, analysis of variance (ANOVA) was performed to confirm the statistical significance of each factor, as shown in Table 10. The optimal parameter combinations were determined by extracting the difference between the extremes of each factor’s mean water value (Ki). The optimal level was determined by selecting the smallest Ki value for the low-optimal indicators.

Through orthogonal experiments and range analysis, this study revealed the discrete effects of three factors, namely, server power combination (A), air supply velocity (B), and server arrangement methods (C), on the thermal performance of servers. As shown in Table 9, air supply velocity (B) played a dominant role in influencing the discretisation degree, with a range of 5.933, which was significantly higher than that of the other factors. Therefore, the importance of the three factors was ranked as B > C > A, i.e., it indicated that optimising the airflow distribution needed to prioritise the adjustment of the server wind speed, followed by the arrangement method, and lastly the power combination. The k—values of the different factors were analysed, in which the least discrete combination of factor A was 400 W × 10, and the difference between the two groups of 500 W × 8 and 1000 W × 4 was not significant. In comparison, the value of the discrete degree of 800 W × 5 was much larger than all three groups. For the B factor, the best discreteness parameter was 4 m/s, and as the wind speed gradient decreased, its k value decreased. Nevertheless, when air supply velocity was 5 m/s, the degree of decrease slightly increased, which might have been due to the closed thermal channel. When the air supply velocity was too high, there was not enough space for the airflow to mix and merge, which affected the airflow’s homogeneity. Finally, the best

	Case7	Case15	Case19	Case23	Case27
Upper vacant area	46.77	39.82	40.35	40.71	40.33
Lower vacant area	31.47	39.32	40.74	25.60	26.02
Temperature differences	15.30	0.5	-0.39	15.11	14.31

Table 7. The temperature difference between the upper and lower vacant areas.

Experimental case	Parameter			Results
	A(W)	B(m/s)	C	Discreteness
13	400 × 10	2	Lower arrangement	4.346
14	400 × 10	3	Upper arrangement	6.890
15	400 × 10	4	Middle arrangement	3.120
16	400 × 10	5	Uniform arrangement	4.032
17	500 × 8	3	Lower arrangement	2.446
18	500 × 8	2	Upper arrangement	10.138
19	500 × 8	5	Middle arrangement	2.522
20	500 × 8	4	Uniform arrangement	4.074
21	800 × 5	4	Lower arrangement	3.914
22	800 × 5	5	Upper arrangement	7.394
23	800 × 5	2	Middle arrangement	12.191
24	800 × 5	3	Uniform arrangement	13.387
25	1000 × 4	5	Lower arrangement	1.600
26	1000 × 4	4	Upper arrangement	3.563
27	1000 × 4	3	Middle arrangement	1.792
28	1000 × 4	2	Uniform arrangement	11.728

Table 8. Orthogonal experimental results.

	Discreteness		
	A	B	C
K1	4.597	9.601	3.076
K2	4.795	6.129	6.996
K3	9.222	3.668	4.906
K4	4.671	3.887	8.306
Range	4.625	5.933	5.229
Significance of factors	B > C > A		
Optimized results	A1 (400 W × 10) B3 (4 m/s) C1 (Lower arrangement)		

Table 9. Analysis of the range of orthogonal experimental results.

Factor	Sum of squares (SS)	Degrees of freedom (df)	Mean square (MS)	F-value	P-value	Significance
A: Server power combination	61.7533	3	20.5844	6.4065	0.0267	*
B: air supply velocity	91.0216	3	30.3405	9.4430	0.0109	*
C: Server arrangement method	63.6949	3	21.2316	6.6080	0.0249	*

Table 10. The analysis results of the *p*-value. Significance: **P* < 0.05; ***P* < 0.01**, **Order of significance: B > C > A.

value of *k* in the C—factor was the lower arrangement with 3.076, followed by the intermediate arrangement. At the same time, the worst degree of discreteness was the homogeneous arrangement with *k* = 8.306.

The ANOVA results (Table 9) further confirm that all three factors have a statistically significant effect (*P* < 0.05), with the order of significance being B > C > A, consistent with the range analysis.

In summary, the orthogonal experiments revealed that the combination of 400 W × 10 power, 4 m/s air velocity, and the lower server arrangement (A1B3C1) was the most promising configuration for minimising the thermal environment's discreteness. While this exact configuration requires dedicated verification experiments, it represents a significant potential improvement over the initial condition and highlights the dominant role of optimising air supply velocity and server arrangement over adjusting the power combination (Fig. 13).

Analysis of actual engineering applications

The main objective of this study is to investigate the impact of internal rack factors on server cooling and propose a least-cost airflow optimisation solution. This approach is designed to operate within the constraints of existing physical and cooling infrastructure, specifically addressing challenges prevalent in edge computing environments and legacy data centre retrofits where structural modifications or adjustments to facility-level

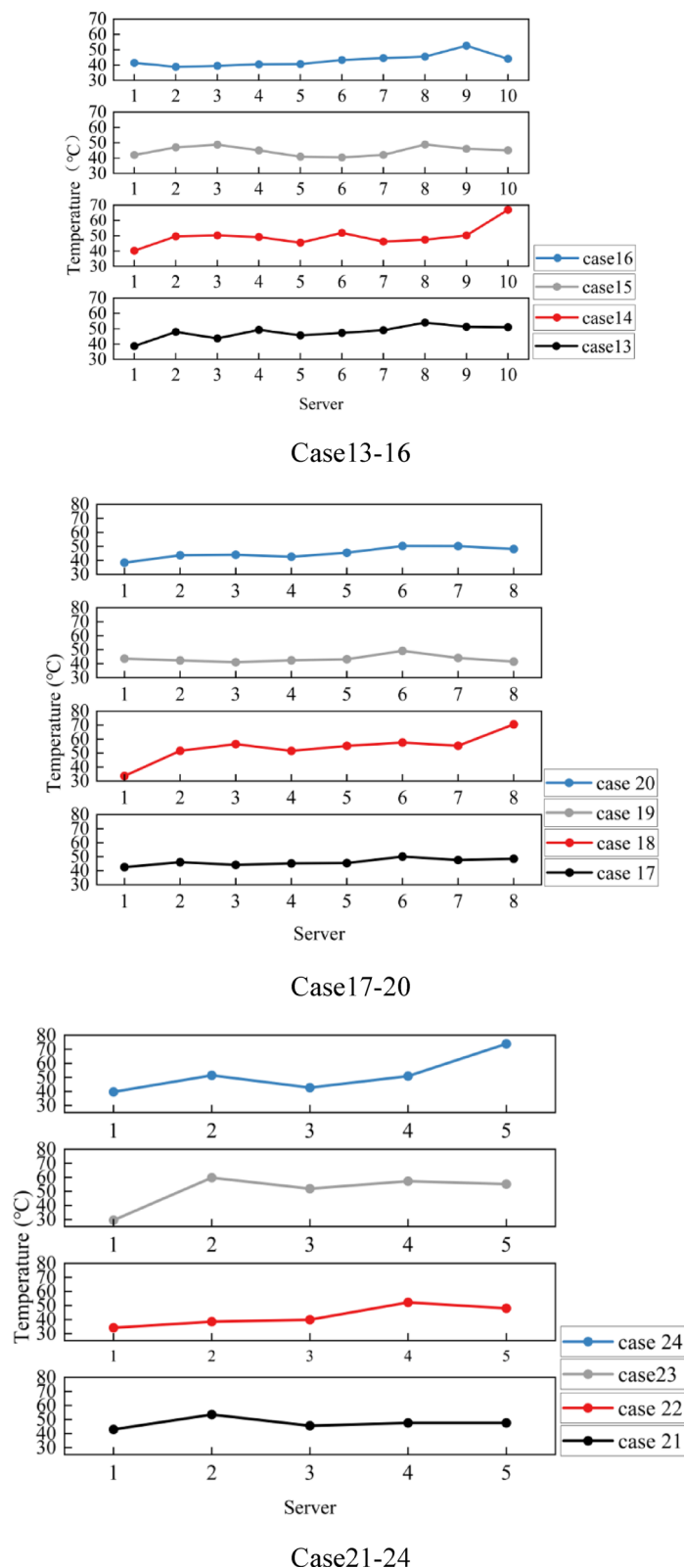
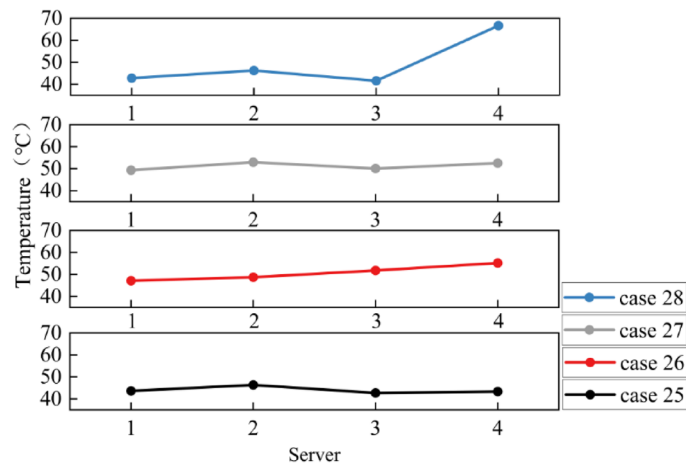


Fig. 13. Orthogonal temperature diagram.

cooling are restricted or impractical. By analysing the individual and combined effects of key parameters, we determine optimal configurations to provide actionable references for energy savings and emissions reduction in established facilities. In a single-factor, four-level experiment, the optimal levels of server power combination, server arrangement method, and air supply velocity are determined to be $400\text{ W} \times 10$, air velocity 4 m/s , and



Case25-28

Fig. 13. (continued)

lower arrangement, respectively. However, due to experimental constraints, this study focuses exclusively on a single-rack configuration. The observed regularities may serve as a useful reference for micro data centres or edge facilities with a small number of racks. For large-scale, multi-rack data centres, caution should be taken when applying these specific conclusions.

Notably, the optimisation scheme proposed in this study offers a distinct economic advantage due to its non-intrusive and low-cost nature. It requires no additional equipment, structural modifications, complex intelligent cooling systems, or adjustments to existing cooling parameters. Instead, the solution involves only minor adjustments to the internal server arrangement within the rack, based on the data centre's existing hardware and infrastructure. As such, it provides a practical, budget-friendly optimisation pathway for data centres with limited financial resources.

There are many types of server power in the data centre, and the higher the server's power, the greater its data processing capacity. If it is convenient to adjust the layout, using servers with higher power and a lower layout is recommended. If it is difficult to change the design, using a server with lower power consumption can reduce the hot-spot temperature.

In this experiment, the larger the air supply velocity, the better the performance of all indicators. However, once the wind speed reaches 4 m/s, the server's cooling efficiency enhancement is no longer apparent, and continued growth will only further increase the economic cost. Therefore, this study suggests that server manufacturers can appropriately increase the server's airflow velocity within a reasonable range.

In a single rack of three-factor, four-level orthogonal experiments to determine the optimal conditions for the internal layout of the rack, it was found that when the power combination of 400 W × 10, the server air supply velocity of 4 m/s, and the arrangement of the lower arrangement of the overall temperature discreteness of the thermal channel of the rack can be the best improvement. Changing the rack's original server power combination will destroy the rack's original operating load, leading to failure of its original air supply system adaptability. Increasing the server air supply speed will result in a slight temperature drop and an increase in energy consumption due to exceeding the air supply system's threshold air speed. Therefore, in the change of internal factors of the rack, when other variables are constant, it is recommended to change the arrangement of the server to the lower part of the arrangement to improve the uniformity of the airflow organisation and then to achieve the overall temperature drop of the system and the homogenization of the thermal field.

Conclusions

This study investigates the influence of rack-server internal arrangement on cooling performance and hot-aisle airflow dynamics, with a focus on the interplay among server power combination, server arrangement method, and air supply velocity. A systematic approach combining single-factor, four-level experiments and orthogonal array testing was employed to assess airflow organisations using discretisation metrics. The analysis elucidates the synergistic effects of multiple variables and their hierarchical impact on rack thermal management, culminating in the identification of an optimal experimental configuration. The findings offer practical insights for enhancing rack heat dissipation in scenarios where structural modifications to the data centre are constrained. Key conclusions are summarised as follows:

1. Increasing server power initially degrades the rack's thermal environment, but beyond an 800 W threshold, the thermal conditions improve significantly due to enhanced inter-rack spacing.
2. Top server outlets create hot spots due to heat buildup, which varies with arrangements. A larger top empty area reduces this effect, making lower arrangements optimal.
3. Under fixed server conditions, higher air supply velocities lower both the average and variation of hot-channel temperatures. However, experimental results indicate that increasing the velocity beyond 4 m/s yields

- only minimal further gains in thermal performance. Thus, 4 m/s offers the best trade-off between cooling performance and energy efficiency.
- Through orthogonal experimental range analysis, the primary and secondary orders of factors affecting the cooling of the server are determined as follows: server air supply wind speed > power configuration > arrangement.
 - The optimal experimental group parameters for the internal arrangement of a single rack are a power combination of 400 W × 10, a server air supply speed of 4 m/s, and the lower arrangement method.

Future research could advance by constructing a multi-rack experimental platform to simulate airflow interference mechanisms between racks in realistic data centre environments, while integrating 3D Computational fluid dynamics (CFD) modelling with MATLAB-based data analysis to quantify how airflow disturbances correlate with server temperature anomalies systematically. These approaches would address limitations in existing single-rack experimental setups and in idealised thermal simulations, thereby improving alignment with real-world operational conditions.

Data availability

All data generated or analysed during this study are included in this published article.

Appendix

Item	Specifications
Rake	Dimensions: 0.6 m × 0.6 m × 2 m Volume: 0.72 m ³
	Quantity: 1
	Unit: 42U
	Total power: 4000 W
Server	Dimensions: 0.44 m × 0.3 m × 0.089 m Volume: 0.012 m ³
	Power levels: 400 W, 500 W, 800 W, 1000 W
Server fan	Dimensions: 0.08 m × 0.08 m Area: 0.0064 m ²
	Airspeeds: 2 m/s, 3 m/s, 4 m/s, 5 m/s
Server arrangements methods	upper, lower, middle, uniform
The ventilated room	Dimensions: 3.6 m × 3.1 m × 3 m Volume: 33.48 m ³
The temperature-controlled room	Dimensions: 3.1 m × 2 m × 3 m Volume: 18.6 m ²
The window connected the ventilated room and the temperature-controlled room	Dimensions: 1.1 m × 0.9 m Area: 0.99 m ²
The snake-shaped heating rod	Dimensions: 0.43 m × 0.15 m

Received: 10 September 2025; Accepted: 16 June 2026
Published online: 25 June 2026

References

- Cho, Y. J., Park, B. R., Hyun, J. Y. & Moon, J. W. Development of an adaptive artificial neural network model and optimal control algorithm for a data center cyber-physical system. *Build. Environ.* **210**, 17. <https://doi.org/10.1016/j.buildenv.2021.108704> (2022).
- Wang, J., Guo, Y., Yue, C., Deng, W. & Zeng, L. Comprehensive assessment of waste heat recovery mismatch and renewable energy integration in data centers: A multifaceted energy, economic, and environmental perspectives. *J. Clean. Prod.* **472**, 13. <https://doi.org/10.1016/j.jclepro.2024.143466> (2024).
- Wang, J., Zhang, Q., Yoon, S. & Yu, Y. Impact of uncertainties on the supervisory control performance of a hybrid cooling system in data center. *Build. Environ.* **148**, 361–371. <https://doi.org/10.1016/j.buildenv.2018.11.026> (2019).
- Tian, Y. et al. Calibration method for sensor drifting bias in data center cooling system using Bayesian Inference coupling with Autoencoder. *J. Build. Eng.* **67**, 14. <https://doi.org/10.1016/j.jobbe.2023.105961> (2023).
- Yang, D. et al. Prosumer data center system construction and synergistic optimization of computing power, electricity and heat from a global perspective. *Therm. Sci. Eng. Progress* **49**, 14. <https://doi.org/10.1016/j.tsep.2024.102469> (2024).
- Zhang, Y., Shan, K., Li, X., Li, H. & Wang, S. Research and technologies for next-generation high-temperature data centers-state-of-the-arts and future perspectives. *Renew. Sustain. Energy Rev.* **171**, 13. <https://doi.org/10.1016/j.rser.2022.112991> (2023).
- Wang, J. et al. Sensor fault diagnosis and correction for data center cooling system using hybrid multi-label random Forest and Bayesian Inference. *Build. Environ.* **249**, 13. <https://doi.org/10.1016/j.buildenv.2023.111124> (2024).
- Du, S., Zhang, Q., Zou, S., Meng, F. & Liu, L. Simulation analysis on energy consumption and economy of CPU cooling system based on loop heat pipe for data center. *Therm. Sci. Eng. Progress* **45**, 15. <https://doi.org/10.1016/j.tsep.2023.102115> (2023).
- MirhoseiniNejad, S., Badawy, G. & Down, D. G. Holistic thermal-aware workload management and infrastructure control for heterogeneous data centers using machine learning. *Future Gener. Comput. Syst. Int. J. eSci.* **118**, 208–218. <https://doi.org/10.1016/j.future.2021.01.007> (2021).
- Zhao, R. et al. A critical review on the thermal management of data center for local hot-spot elimination. *Energy Build.* **297**, 16. <https://doi.org/10.1016/j.enbuild.2023.113486> (2023).
- Arumugam, J., Maiya, M. P. & Nagendra, S. M. S. Experimental evaluation of thermal characteristics of wall mounted radiant cooled environment with ceiling fan. *Sadhana-Academy Proc. Eng. Sci.* **49**, 14. <https://doi.org/10.1007/s12046-024-02480-5> (2024).

12. Tatchell-Evans, M., Kapur, N., Summers, J., Thompson, H. & Oldham, D. An experimental and theoretical investigation of the extent of bypass air within data centres employing aisle containment, and its impact on power consumption. *Appl. Energy* **186**, 457–469. <https://doi.org/10.1016/j.apenergy.2016.03.076> (2017).
13. Chu, W., Wang, R., Hsu, P. & Wang, C. Assessment on rack intake flowrate uniformity of data center with cold aisle containment configuration. *J. Build. Eng.* **30**, 10. <https://doi.org/10.1016/j.job.2020.101331> (2020).
14. Zhan, B. et al. Experimental investigation on ducted hot aisle containment system for racks cooling of data center. *Int. J. Refrig* **127**, 137–147. <https://doi.org/10.1016/j.ijrefrig.2021.02.006> (2021).
15. Tradat, M. I., Manaserh, Y. M. A., Sammakia, B. G., Hoang, C. H. & Alissa, H. A. An experimental and numerical investigation of novel solution for energy management enhancement in data centers using underfloor plenum porous obstructions. *Appl. Energy* **289**, 18. <https://doi.org/10.1016/j.apenergy.2021.116663> (2021).
16. Lu, H., Zhang, Z. & Yang, L. A review on airflow distribution and management in data center. *Energy Build.* **179**, 264–277. <https://doi.org/10.1016/j.enbuild.2018.08.050> (2018).
17. Wu, L. et al. An experimental study on improving the thermal environment of data centers by optimizing fan wall arrangement. *J. Build. Eng.* **96**, 17. <https://doi.org/10.1016/j.job.2024.110421> (2024).
18. Zhang, J. et al. Optimization air-conditioning system and thermal management of data center via fan-wall free cooling technology. *Appl. Therm. Eng.* **234**, 12. <https://doi.org/10.1016/j.applthermaleng.2023.121245> (2023).
19. Wang, W. L., Li, Y. & Lee, K. P. Effects of airflow distribution changes on the thermal environment of IT equipment in data centers. *Case Stud. Therm. Eng.* **73**, 106604. <https://doi.org/10.1016/j.csite.2025.106604> (2025).
20. Omrani, M. & Ghassemi, M. AI-driven optimization of fan-wall cooling system in a medium-density data center. *Int. J. Heat Mass Transf.* **247**, 127159. <https://doi.org/10.1016/j.ijheatmasstransfer.2025.127159> (2025).
21. Chu, W. & Wang, C. A review on airflow management in data centers. *Appl. Energy* **240**, 84–119. <https://doi.org/10.1016/j.apenergy.2019.02.041> (2019).
22. Li, D., Zhang, Y., Song, J., Liu, H. & Jiang, J. Energy saving with zero hot spots: A novel power control approach for sustainable and stable data centers. *Sustainability* **14**, 17. <https://doi.org/10.3390/su14159005> (2022).
23. Capozzoli, A. & Primiceri, G. Cooling systems in data centers: State of art and emerging technologies. *Energy Proc.* **83**, 484–493. <https://doi.org/10.1016/j.egypro.2015.12.168> (2015).
24. Wu, L. et al. Optimization of airflow organization in fan-wall data center via baffles. *Appl. Therm. Eng.* **244**, 12. <https://doi.org/10.1016/j.applthermaleng.2024.122745> (2024).
25. Kuzay, M. et al. Retrofitting of an air-cooled data center for energy efficiency. *Case Stud. Therm. Eng.* **36**, 14. <https://doi.org/10.1016/j.csite.2022.102228> (2022).
26. Lim, S. & Chang, H. Airflow management analysis to suppress data center hot spots. *Build. Environ.* **197**, 19. <https://doi.org/10.1016/j.buildenv.2021.107843> (2021).
27. Yuan, X. et al. Design and validation of an airflow management system in data center with tilted server placement. *Appl. Therm. Eng.* **164**, 12. <https://doi.org/10.1016/j.applthermaleng.2019.114444> (2020).
28. Yuan, X., Wang, Y., Liu, J., Xu, X. & Yuan, X. Experimental and numerical study of airflow distribution optimisation in high-density data centre with flexible baffles. *Build. Environ.* **140**, 128–139. <https://doi.org/10.1016/j.buildenv.2018.05.043> (2018).
29. Yuan, X. et al. Experimental and numerical investigation of an airflow management system in data center with lower-side terminal baffles for servers. *Build. Environ.* **155**, 308–319. <https://doi.org/10.1016/j.buildenv.2019.03.039> (2019).
30. Yuan, X., Liu, J., Yang, Y., Wang, Y. & Yuan, X. Investigation and improvement of air distribution system's airflow performance in data centers. *Proc. Eng.* **205**, 2895–2902. <https://doi.org/10.1016/j.proeng.2017.09.920> (2017).
31. Jiao, J., Zhu, C., Li, Y. & Gong, L. A numerical investigation of enhancing supply air flowrate uniformity for data centre with in-row cooling system. *J. Build. Eng.* **121**, 115656. <https://doi.org/10.1016/j.job.2026.115656> (2026).
32. Lan, J. et al. Experimental and numerical investigation on thermal performance of data center via fan-wall free cooling technology. *Appl. Therm. Eng.* **228**, 14. <https://doi.org/10.1016/j.applthermaleng.2023.120467> (2023).
33. Song, P., Zhang, Z. & Zhu, Y. Numerical and experimental investigation of thermal performance in data center with different deflectors for cold aisle containment. *Build. Environ.* **200**, 107961. <https://doi.org/10.1016/j.buildenv.2021.107961> (2021).
34. Lenovo, Thermal rules—ThinkSystem SR630 V2, Lenovo Documentation, https://pubs.lenovo.com/sr630-v2/zh-CN/thermal_rules (2024).
35. Sharma, R., Bash, C. & Patel, C. Dimensionless parameters for evaluation of thermal design and performance of large-scale data centers. *Am. Inst. Aeronaut. Astronaut.* <https://doi.org/10.2514/6.2002-3091> (2002).
36. Zhang, K., Zhang, Y., Liu, J. & Niu, X. Recent advancements on thermal management and evaluation for data centers. *Appl. Therm. Eng.* **142**, 215–231. <https://doi.org/10.1016/j.applthermaleng.2018.07.004> (2018).
37. Cho, J., Yang, J. & Park, W. Evaluation of air distribution system's airflow performance for cooling energy savings in high-density data centers. *Energy Build.* **68**, 270–279. <https://doi.org/10.1016/j.enbuild.2013.09.013> (2014).
38. Wang, L., Tang, Y., Bai, X., Mi, R. & Wang, Y. Experimental and optimization research of the rack thermal environment based on the dynamic server power. *J. Build. Eng.* **98**, 9. <https://doi.org/10.1016/j.job.2024.111314> (2024).
39. Pecht, M., Lall, P. & Hakim, E. The influence of temperature on integrated circuit failure mechanisms. *Qual. Reliab. Eng. Int.* **11**, 399–411. <https://doi.org/10.1002/qre.4680110608> (1995).
40. Moffat, R. J. Describing the uncertainties in experimental results. *Exp. Thermal Fluid Sci.* **1**(1), 3–17. [https://doi.org/10.1016/0894-1777\(88\)90043-X](https://doi.org/10.1016/0894-1777(88)90043-X) (1988).
41. Li, C. et al. Experimental study on direct evaporative cooling for free cooling of data centers. *Energy* **288**, 13. <https://doi.org/10.1016/j.energy.2023.129889> (2024).
42. The Green Grid, Green Grid metrics: Describing data center power efficiency, The Green Grid White Paper #1, 8. https://www.thegreengrid.org/~media/WhitePapers/White_Paper_1_Data_Center_Power_Efficiency_Metrics_2007.pdf (2007).

Author contributions

G: Writing—review and editing, Writing—original draft, Formal analysis, Data curation, Conceptualization. C: Writing—review and editing, Methodology, M: Supervision, Resources, Methodology, Conceptualization, Funding acquisition. H: Supervision, Formal analysis. W: Supervision, Resources. ZQ: Supervision, Methodology, Conceptualization. Y: Writing—review and editing. ZZ: Validation, Funding acquisition. Y: Supervision.

Funding

The author would like to thank the financial support of the Guizhou University Laboratory Open Project (No. SYSKF2025-026 and No. SYSKF2026-017), Guizhou Province Postgraduate Research Fund Project (No.2024YJSKYJ114), Undergraduate Innovation and Entrepreneurship Training Program Project (No. gzusc2025127) and Guizhou Provincial Basic Research Program (Natural Science) (Qiankehe Foundation MS[2026]181).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to R.M. or C.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026