Sustainable Energy Technologies and Assessments 82 (2025) 104511

Contents lists available at ScienceDirect

Sustainable Energy Technologies and Assessments

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

Original article

AI-driven cooling technologies for high-performance data centres: state-of-the-art review and future directions☆

Dlzar Al Kez [a,b,c,*], Aoife M Foley [a,b,c], Fadhli Wong B.M. Hasan Wong [d], Andrea Dolfi [e], Geetha Srinivasan [e,f]

[a] Department of Civil Engineering and Management, School of Engineering, The University of Manchester, Manchester, the United Kingdom of Great Britain and Northern Ireland
[b] Tyndall Centre for Climate Change Research, School of Engineering, The University of Manchester, Manchester, the United Kingdom of Great Britain and Northern Ireland
[c] Department of Electrical & Electronic Engineering, School of Engineering, The University of Manchester, Manchester, the United Kingdom of Great Britain and Northern Ireland
[d] School of Mechanical and Aerospace Engineering, Queen's University of Belfast, Belfast, the United Kingdom of Great Britain and Northern Ireland
[e] Group Research and Technology, PETRONAS, Kawasan Institusi Bangi, Malaysia
[f] School of Chemistry and Chemical Engineering, Queen's University of Belfast, the United Kingdom of Great Britain and Northern Ireland

ARTICLE INFO

Keywords:
AI-driven cooling
Commercial readiness
Data centres
Energy efficiency
Hybrid cooling
Immersion cooling
Liquid cooling
Thermal management

ABSTRACT

The increasing computational demands of artificial intelligence (AI), high-performance computing (HPC), and hyperscale cloud platforms are placing significant thermal and energy pressures on data centre infrastructure. Traditional air-based cooling systems are increasingly inadequate for managing these loads, prompting a transition toward more efficient, scalable, and sustainable alternatives. This study presents a comprehensive, system-wide review of next-generation cooling technologies, including direct liquid cooling, immersion cooling, two-phase systems, spray and jet impingement cooling, and heat pipe-based solutions. Unlike previous reviews focused on component-level or single-technology evaluations, this study integrates technical performance, commercial readiness, and environmental impact across diverse deployment conditions. A detailed comparative framework synthesises thermal efficiency, scalability, and water usage across air, liquid, and hybrid systems. Special attention is given to commercially mature solutions such as RDHx and cold plate DLC, while the feasibility of emerging methods like AI-driven cooling, phase-change materials, and thermoelectric technologies is evaluated. The review further explores heat reuse potential and ESG-aligned design strategies critical to decarbonising digital infrastructure. By mapping trade-offs across performance, cost, and sustainability, this study offers actionable insights for data centre operators, designers, and policy stakeholders navigating the path to high-efficiency, AI-ready cooling.

Introduction

The global digital economy is expanding at an unprecedented rate, driven by cloud computing, artificial intelligence (AI), 5G networks, and the Internet of Things (IoT). As industries digitise and computational workloads grow in complexity, data centres have become the backbone of modern digital infrastructure, enabling real-time data processing, storage, and transmission. In 2023, the global data centre market was valued at approximately $219.23 billion, with projections indicating a compound annual growth rate (CAGR) of 11.6 %, reaching $584.86 billion by 2032 [1]. As computational power requirements increase, so does the demand for efficient cooling solutions to sustain high-performance operations. This rapid expansion is particularly evident in hyperscale data centres, operated by cloud service giants (i.e., Amazon Web Services (AWS), Microsoft Azure, and Google Cloud), with over 992 hyperscale facilities globally as of 2023 [2].

While this rapid expansion fuels digital transformation, it also raises critical challenges related to energy consumption and thermal management [3]. Data centres already account for 1–1.5 % of global electricity use, and with the rise of AI-driven workloads, energy demand is

---

☆ This article is part of a special issue entitled: 'SDEWES 2024 Conference' published in Sustainable Energy Technologies and Assessments.

* Corresponding author at: Department of Civil Engineering and Management, School of Engineering, The University of Manchester, Manchester, United Kingdom.
E-mail addresses: dalkez01@qub.ac.uk, dlzar.alkez@manchester.ac.uk (D. Al Kez).

https://doi.org/10.1016/j.seta.2025.104511
Received 3 February 2025; Received in revised form 22 July 2025; Accepted 6 August 2025

projected to double by 2026 [4,5]. AI models (i.e., ChatGPT and DeepSeek) require exponentially increasing computational power, with each query consuming 2.9 Wh of electricity, ten times higher than a traditional search engine request [3,6]. If AI-generated responses were integrated into global search engines, per-query energy consumption could rise to 6.9–8.9 Wh, significantly impacting power demand. Although newer models like DeepSeek are being designed to be more energy-efficient, the sheer volume of AI applications and the accelerating pace of digitalisation continue to drive an unprecedented rise in computational energy demand. Moreover, the actual power consumption of most data centres remains opaque, as energy reporting varies across operators and regions, but estimates suggest the global footprint is far higher than officially disclosed, with energy-intensive AI workloads further amplifying the challenge [5]. This escalating energy footprint underscores the urgent need for next-generation cooling strategies that optimise energy efficiency while ensuring computational sustainability.

For instance, Microsoft is partnering with renewable energy developers to add capacity for their data centres in the U.S. and Europe [7]. Similarly, Digital Realty, a major data centre provider, has committed to powering its facilities with 100 % renewable energy, achieving this goal across its European portfolio and U.S. colocation data centres [8]. Despite these efforts, the increasing energy consumption of data centres, particularly due to AI applications, presents challenges. While the transition to renewable energy sources can help offset emissions, renewables alone cannot fully address the cooling efficiency problem. The sheer growth in data centre energy demand necessitates further advancements in sustainable and adaptive cooling technologies to complement sustainability efforts.

A substantial portion (40 %) of data centre energy consumption is attributed to cooling systems, which are crucial for maintaining hardware reliability and operational efficiency [9]. Traditionally, air-based cooling solutions (i.e., Computer Room Air Conditioning (CRAC) and Computer Room Air Handling (CRAH) units) have been employed due to their simplicity and cost-effectiveness [10]. However, as data centre power densities increase, these systems are struggling to keep pace, leading to hotspots, inefficiencies, and reduced equipment lifespans [11,12]. Although hot and cold aisle containment (CAC) strategies have helped improve cooling efficiency, they remain insufficient for the thermal demands of modern AI models and high-performance computing (HPC) workloads [13]. Consequently, the industry is shifting toward liquid-based and AI-enhanced cooling architectures to bridge the efficiency gap.

Emerging cooling solutions such as direct liquid cooling (DLC), immersion cooling, phase-change cooling, and AI-driven predictive thermal management are increasingly being explored to address these challenges [14]. Liquid cooling has emerged as a viable alternative due to its superior heat dissipation capabilities and lower energy consumption compared to air-based methods [15]. Liquids possess significantly higher thermal conductivity and volumetric heat capacity than air, allowing for more efficient heat removal and a potential reduction in total cooling-related energy costs [16,17]. Direct-to-chip cooling, which places cold plates directly on central processing units (CPUs) and graphics processing units (GPUs), achieves high heat capture ratios (∼94 %), offering substantial cooling energy savings [18]. However, implementation challenges (i.e., fluid leakage risks, infrastructure costs, and legacy system compatibility) hinder its widespread adoption [19].

Immersion cooling, a more disruptive alternative, submerges server components in thermally conductive dielectric fluids, achieving heat transfer rates up to 1,000 times greater than air cooling [20,21]. Despite its efficiency, widespread adoption remains limited due to economic and technical barriers [22]. The high initial capital expenditure associated with immersion cooling infrastructure (i.e., customised racks, specialised dielectric fluids, and sealed enclosures) poses a significant challenge. Additionally, prolonged exposure to dielectric fluids can lead to material compatibility issues, including swelling, degradation, and corrosion of electronic components, necessitating specialised coatings and modifications [23]. Furthermore, the degradation of certain dielectric fluids over time and the accumulation of contaminants can negatively impact their thermal performance, requiring periodic filtration or replacement [24].

Several reviews have examined data centre cooling technologies, each providing valuable insights into different aspects of thermal management. Pambudi et al. conducted a comprehensive review of immersion cooling technologies, emphasising their historical evolution, energy savings potential, and associated challenges (i.e., dielectric fluid degradation and maintenance complexities) [25]. Zhou et al. explored mechanical refrigeration and hybrid cooling systems, focusing on the role of free cooling technologies such as heat pipes and evaporative cooling in reducing Power Usage Effectiveness (PUE) [26]. It is important to note that PUE, as defined by the ISO/IEC 30134–2 standard, is not a measure of total system efficiency [27]. Rather, it is a ratio of total facility energy to the energy used by IT equipment, offering insight into non-ICT energy overhead such as cooling, power distribution, and lighting. Despite its name, PUE does not capture the full energy efficiency of the data centre. Li et al. examined non-uniform load distribution in data centres, proposing microchannel liquid cooling as an effective solution for improving heat dissipation and reducing thermal hotspots [28]. While their review highlighted the advantages of chip-scale thermal management, it left room for broader evaluations of full-system cooling architectures.

Other reviews have examined the role of AI in data centre cooling. Zhou et al. identified AI-based control systems as an emerging research direction capable of optimising server-level cooling efficiency in real-time [26]. However, their study did not provide a detailed comparison of immersion and liquid cooling strategies at a large scale. Azarifar et al. investigated direct-to-chip and immersion cooling strategies, focusing on their feasibility in hyperscale and edge data centres [29]. While they identified manifold liquid cooling as a promising technology for high-power-density racks, their study lacked a comprehensive evaluation of implementation challenges, cost trade-offs, and long-term sustainability. Ziyong Li et al. provided an extensive review of chip-scale thermal management, covering air cooling, liquid cooling, and emerging electrocaloric cooling techniques [30]. They underscored the importance of multi-phase liquid cooling for AI-centric workloads but focused primarily on chip-level solutions rather than whole-system approaches.

Despite the insights provided by existing studies, a comprehensive, system-wide analysis of advanced cooling architectures, particularly their feasibility for AI-driven and high-performance data centres, remains missing. This paper aims to fill that gap by critically evaluating these next-generation cooling technologies, assessing their long-term viability, and identifying practical implementation challenges and opportunities for industry adoption. Additionally, this review explores the role of AI-enhanced cooling strategies, an area that remains understudied in current literature, to provide a forward-looking perspective on next-generation thermal management solutions.

To address this gap, this study adopts a system-wide perspective, analysing cooling strategies through the lens of deployment readiness, commercial scalability, regional resource constraints, and AI-enhanced thermal control. In addition to evaluating technical performance, this review provides a comparative assessment of hybrid cooling pathways, explores AI integration across facility-wide thermal architectures [31], and considers emerging materials and smart fluids in the context of Environmental, Social, and Governance (ESG) alignment and operational resilience [32]. This approach allows for a more holistic understanding of next-generation data centre cooling that bridges academic innovation with real-world implementation.

This paper contributes to the field by synthesising recent advancements in data centre cooling technologies and evaluating their feasibility, limitations, and potential for integration into high-density computing environments. By adopting a comparative and application-focused approach, this study provides actionable insights for researchers, engineers, and industry professionals, bridging the gap between theoretical advancements and real-world deployment.

Research framework

This study adopts a hybrid narrative-systematic review approach to evaluate next-generation data centre cooling technologies. The objective is to move beyond isolated or chip-level reviews by combining a structured literature screening process with comparative analysis of technical performance, deployment maturity, and sustainability dimensions.

A systematic literature review was conducted using targeted keywords including: “direct liquid cooling,” “immersion cooling,” “spray cooling,” “jet impingement cooling,” “microchannel cooling,” “manifold microchannel cooling,” “two-phase cooling,” “phase-change cooling,” “thermoelectric cooling,” “dielectric fluids,” “data centre cooling,” “high-density computing,” and “AI-driven cooling.” Searches covered peer-reviewed publications indexed in Web of Science, Scopus, IEEE Xplore, and ScienceDirect, spanning the period from 2015 to 2025. This timeframe captures the rapid emergence of AI-accelerated workloads, increasing power densities, and growing adoption of liquid-based and hybrid thermal solutions.

In addition to academic literature, the review incorporates grey literature sources such as white papers, vendor case studies, manufacturer blogs, technical specifications, and government or institutional reports (e.g., International Energy Agency (IEA), U.S. Department of Energy (DOE), Uptime Institute, ASHRAE). These sources were used selectively to validate real-world deployments, extract design data (e.g., rack power densities, heat reuse, water usage), and benchmark commercial readiness.

To ensure relevance and technical depth, studies were screened for inclusion based on the following criteria: (i) discussion of architecture-level cooling systems (not limited to chip-scale); (ii) provision of empirical or modelled data (e.g., PUE, Water Usage Effectiveness (WUE), outlet temperatures, supported rack densities); and (iii) relevance to AI/HPC thermal demands or large-scale deployment feasibility. Publications were excluded if they were non-English, lacked performance data, or focused only on niche/experimental materials without architectural integration. The selected materials were categorised and analysed across key dimensions: 1) Thermal performance and cooling efficiency; 2) Water use and sustainability trade-offs; 3) Scalability and integration feasibility; 4) Commercial availability and deployment readiness; and 5) Heat reuse and ESG alignment. These dimensions directly informed the structuring of the technical review (Section 4), the comparative analysis of commercial solutions (Section 5), and the synthesis of emerging innovations (Section 6).

To visualise the data flow, Fig. 1 presents the research framework used in this study. It maps the process from source identification to screening and analytical integration, showing how three data streams, peer-reviewed literature, industry reports, and web-based technical content, were processed to support a system-wide comparative evaluation. This engine underpins the benchmarking tables and strategic insights presented in later sections.

The remainder of this paper is structured as follows: Section 2 reviews traditional cooling methods; Section 3 discusses advancements in air-based advanced strategies; Section 4 explores liquid cooling data centre solutions; Section 5 examines commercial readiness and deployment trends across air, liquid, and hybrid systems. Section 6 outlines future trends and research directions. Finally, Section 7 concludes with key takeaways and recommendations.

Traditional cooling systems

Computer room air conditioning

Traditional CRAC systems have long been the cornerstone of data centre cooling, using air- or refrigerant-based methods to maintain optimal server room temperatures. As shown in Fig. 2, these systems circulate cold air for recirculation using chilled water loops, cooling towers, and pumps.

Their ease of implementation and low initial cost made them ideal for small to medium-sized data centres. However, they are increasingly inadequate for modern high-density, AI-driven workloads due to their

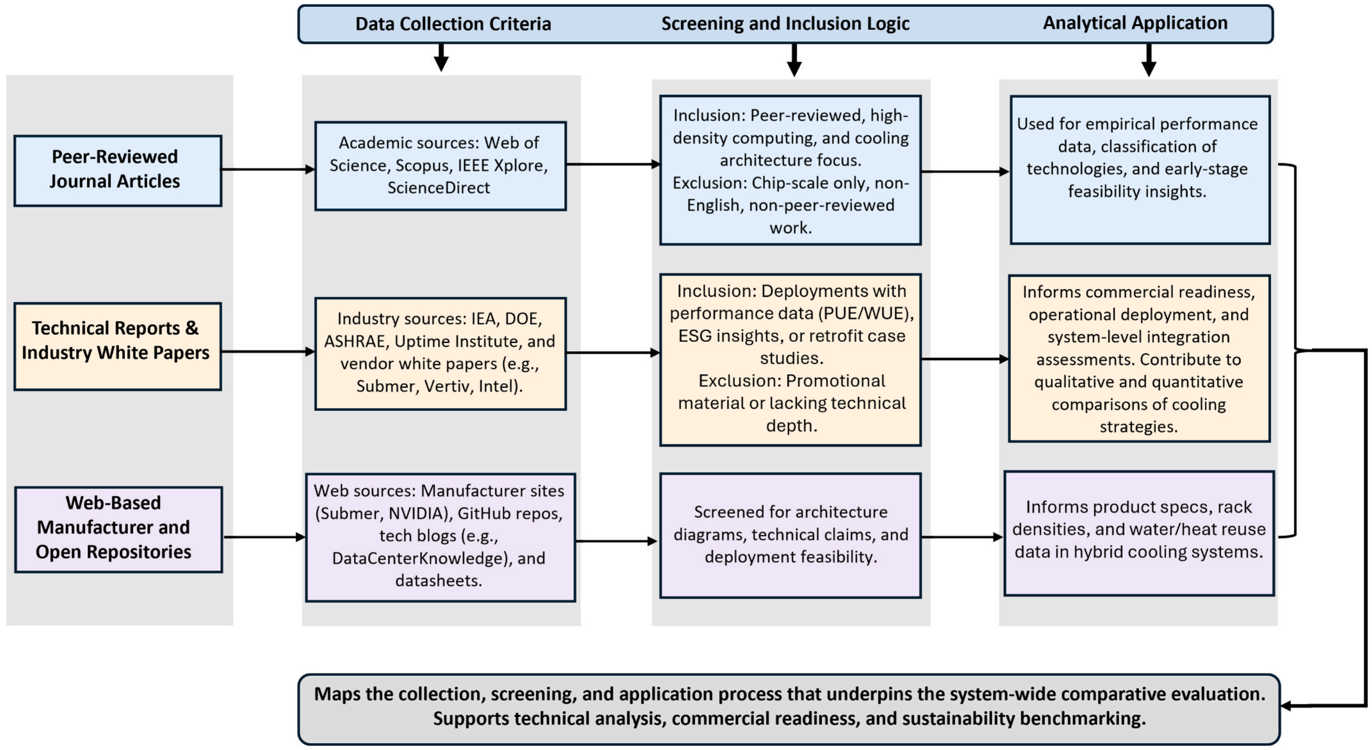

Fig. 1. Systematic research framework showing source types, screening criteria, and analytical applications used to construct the comparative analysis engine.

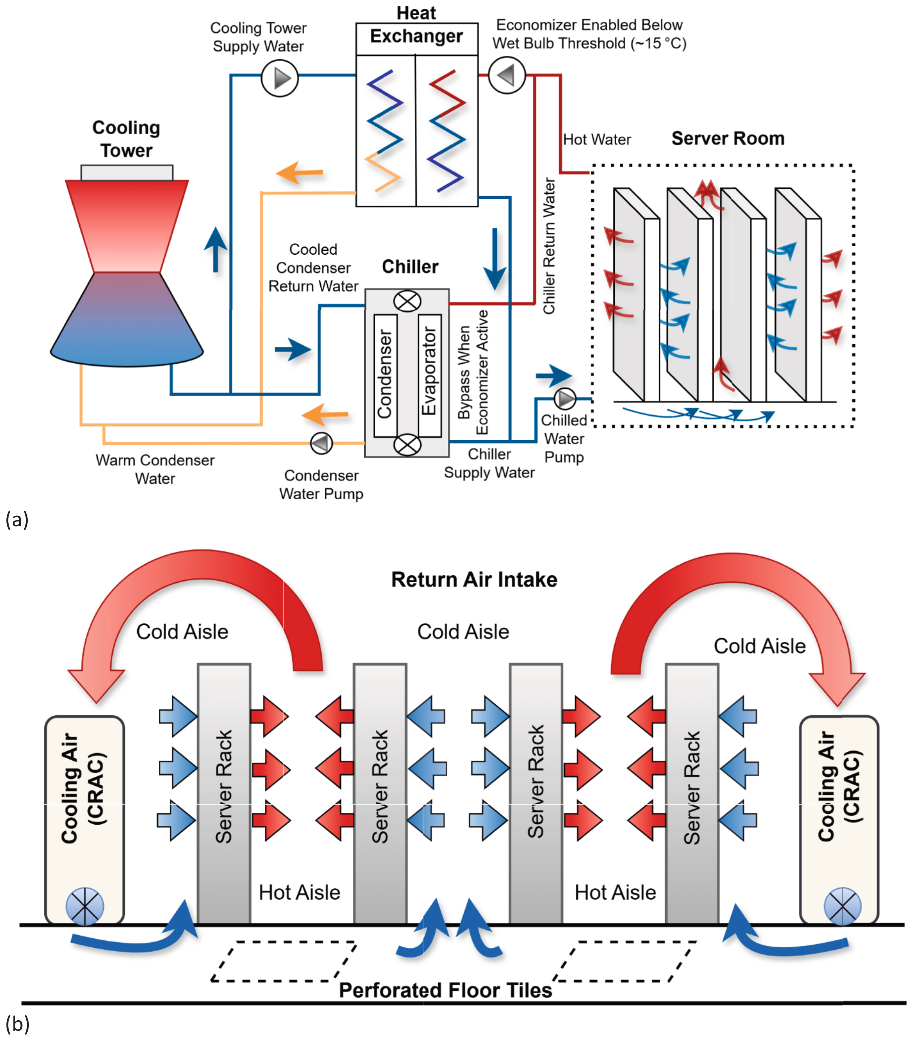

Fig. 2. (a) Schematic of a traditional CRAC-based cooling system, air-side economiser, and water-side economiser. Adapted from [33]. (b) Schematic of a hot/cold aisle containment approach. Source: Authors.

dependency on ambient conditions [34]. For example, elevated outdoor temperatures reduce heat exchange efficiency, increasing compressor pressures and energy consumption. While typical coefficient of performance (COP) values range from 3.14 to 3.17, they decline in extreme climates [35]. CRAC systems often rely on constant air volume (CAV) designs, which lack the flexibility to adapt to variable thermal demands, resulting in localised overcooling or inadequate cooling [36]. Inefficiencies like hot air recirculation and bypass airflow can reduce system performance by up to 20 % [37]. As heat dissipation from AI workloads grows less predictable, airflow management strategies (i.e., hot/cold aisle containment and dynamic control) are increasingly explored to address these issues.

Geometric and operational parameters significantly affect CRAC performance. For example, plenum heights of 0.6 m improve airflow uniformity and reduce temperature differentials by over 20 % [38]. Inline CRAC arrangements also enhance rack temperature balance compared to perpendicular layouts. However, traditional systems remain energy-intensive under fluctuating AI loads. Hybrid setups combining CRAC with liquid cooling have shown promise, achieving COPs as high as 15.9 and efficiency improvements over 30 % [39]. Despite limitations, CRAC and CRAH systems continue to dominate commercially, with the global market projected to reach USD 31.95 billion by 2030 at a CAGR of 13.9 %, driven by legacy use and modular retrofit potential [40].

Airflow distribution inefficiencies due to leakage and bypass airflow (LBA) further hinder CRAC performance. One study using a 2D airflow-heat transfer model optimised the supply rate, reducing racks with a Supply Heat Index (SHI) greater than 0.2 from 39.5 % to 17.2 %, saving 406.24 kW in energy compared to overcooling solutions [41]. However, accurately assessing energy savings remains difficult due to workload and climate variability.

While CRAC systems continue to support legacy data centres, their reliance on mechanical refrigeration, high energy consumption, and limited adaptability to variable thermal loads increasingly constrain their suitability for modern needs. This is especially true in AI-intensive environments, where heat densities are rapidly escalating. Liquid cooling is emerging as a more scalable and efficient alternative, offering superior performance for high-density workloads. To bridge this gap, advanced modelling tools like Computational Fluid Dynamics (CFD) integrated with Building Energy Modelling (BEM) have shown promise. These models better capture non-uniform airflow and temperature distributions, enabling improved system design and control [42]. However, their reliance on precomputed parameters limits real-time applicability, particularly under the fluctuating loads typical of AI workloads.

As such, there is a growing need for intelligent, adaptive thermal management systems that respond dynamically to workload and climate variability. The future of sustainable data centre cooling will likely centre on a shift toward liquid-based systems, enhanced airflow control, and AI-driven optimisations that can keep pace with evolving computational demands.

Computer room air handler

CRAH systems are central to traditional data centre cooling, using chilled water loops and air-to-water heat exchangers to maintain Information Technology Equipment (ITE) thermal conditions [43]. Compared to refrigerant-based CRAC units, CRAH systems offer better scalability but involve greater infrastructure complexity. They remain widely adapted but face growing limitations as AI-driven and HPC workloads intensify cooling demands.

One key drawback is high energy consumption, CRAH systems can account for 30–50 % of total data centre energy use [44]. Inefficiencies from non-uniform airflow and leakage/LBA can cause hotspots and overcooling [45]. Design optimisations like plenum height and perforated tile configurations significantly impact performance. A laboratory study found that 50 % of porous tiles delivered optimal airflow uniformity under partial containment, improving supply airflow usage by 5 % and reducing pressure drop by 67 % compared to 32 % of tiles [46]. However, even with such optimisations, CRAH systems are difficult to scale for hyperscale data centres due to the significant infrastructure and high water requirements of chilled water distribution networks.

Advanced bypass (BP) methods enhance CRAH efficiency by reducing fan energy consumption. Induced bypass (iBP) creates low-pressure zones and typically outperforms forced bypass (fBP) in airflow uniformity and energy savings [47]. Integrating BP systems with optimised underfloor airflow layouts reduces hotspots and improves cooling [48]. When combined with economisers, BP methods can yield significant energy savings. However, real-time control remains a challenge in AI-driven environments with fluctuating heat loads.

Modern CRAH systems incorporate Variable Speed Drives (VSDs) and smart control to adjust airflow based on real-time CPU loads, improving part-load efficiency [42]. Integrating water-side economisers, which use cool outdoor air to pre-chill water, can cut energy use by up to 47 % annually compared to air-cooled systems [49]. Optimised plenum configurations and containment strategies further reduce bypass airflow, enhance thermal uniformity, and minimise energy consumption in legacy data centres [50]. Despite improvements, quantifying CRAH energy savings is difficult due to operational variability. Additionally, their long-term viability for AI-centric data centres is uncertain, as liquid cooling solutions provide more efficient and scalable thermal management.

Hot/cold aisle containment

Hot and CAC strategies are key to optimising airflow management and enhancing energy efficiency in data centres. These methods involve physically separating cold supply air from hot exhaust air using barriers like panels or doors, minimising air recirculation and bypass. This separation ensures stable inlet temperatures for server racks, reduces cooling workloads, and improves energy efficiency [51]. Server racks are organised into alternating rows of hot and cold aisles, with cold aisles facing air conditioners and hot aisles directing exhaust air away. As AI-driven computing and high-performance workloads generate increasing heat densities, the effectiveness of containment strategies becomes even more crucial in mitigating thermal imbalances and optimising energy efficiency. Fig. 2b illustrates a typical hot/cold aisle containment layout using CRAC units and a raised floor cooling design [52]. Containment strategies, therefore, enhance thermal performance, reduce energy use, and support sustainability goals.

Experimental studies validate the energy-saving potential of containment systems. For instance, a study comparing a contained hot aisle with an uncontained setup found a 25 % reduction in cooling energy consumption, attributed to minimised hot air recirculation and cold air bypass [53]. Similarly, CAC improved rack intake temperatures and achieved near-ideal Rack Cooling Index (RCI) values under high cooling loads. However, incomplete containment led to localised hotspots due to air leakage, highlighting the importance of precise design and installation [54]. For hyperscale and AI-focused data centres, scaling containment solutions across thousands of racks presents logistical challenges, requiring dynamic control mechanisms and real-time airflow optimisation to maintain efficiency.

CFD analyses further demonstrate the advantages of containment strategies. Hot aisle containment (HAC) was shown to reduce recirculation airflow, achieving a 7 % improvement in Return Temperature Index (RTI) and an 8 % increase in Return Heat Index (RHI) [55]. Additionally, installing blanking panels in unoccupied racks minimises bypass airflow, enhancing cooling performance and energy efficiency. CFD simulations also revealed that HAC combined with raised floors improved air distribution efficiency by 28 % and reduced recirculation rates by 40 % compared to hard-floor configurations under typical cooling loads [56]. While containment strategies optimise airflow efficiency within traditional air-cooled data centres, their effectiveness may be limited in high-density AI clusters, where direct-to-chip or immersion cooling offers superior heat dissipation.

In one study, CAC with underfloor precision air conditioning improved Air Supply Efficiency (ASE) from 65.69 % to 85.57 % and reduced the SHI from 0.141 to 0.0027 [57]. HAC showed similar benefits, increasing ASE to 90.25 % and reducing SHI to 0.0024. These results underscore the effectiveness of both CAC and HAC in minimising temperature gradients at cabinet inlets, reducing hotspots, and achieving uniform cooling. HAC delivered the best overall performance in optimising thermal environments and energy use. However, despite controlled experimental validations, real-world efficiency gains from containment remain difficult to quantify due to variations in rack densities, IT load profiles, and cooling infrastructure heterogeneity.

Refinements in containment design contribute significantly to improving cooling efficiency and effectiveness. For example, introducing a triangular deflector below air supply grilles eliminated flow reversal and improved intake flow rate distribution across racks. The deflector reduced the Flowrate Uniformity Index (FUI) standard deviation by 51.3–62.0 %, ensuring uniform airflow [58]. Field studies also demonstrated that enclosed aisle designs minimise bypass airflow, reducing energy consumption by up to 43 % compared to open layouts while maintaining acceptable server temperatures [59].

Despite these benefits, several critical challenges persist, particularly in ensuring airtight containment and adaptive thermal control across dynamic IT workloads. Incomplete containment or poorly designed airflow management can lead to air leakage, localised hotspots, and uneven cooling. Addressing these issues requires advanced sealing techniques, precise installation, and dynamic control systems to adapt to variable IT loads. Furthermore, as AI-driven workloads demand increasing thermal management efficiency, containment strategies alone may not be sufficient, requiring integration with liquid cooling systems or AI-based airflow optimisation to achieve sustainable and scalable cooling solutions. When implemented correctly, containment systems represent a critical component in achieving energy-efficient and sustainable data centre operations.

Air-based advanced cooling

Free air cooling

Free air cooling is an effective method for improving data centre energy efficiency by leveraging natural environmental conditions. It operates in two modes: direct, where untreated external air enters the server room, and indirect, where external air cools indoor air via air-to-air or air-to-water heat exchangers. Depending on location and climate, optimised strategies can reduce energy consumption by 47.5 % to 80 % [60]. Free cooling also supports low PUE and WUE values, reinforcing its role in sustainable operations.

Air-side economisation is central to free cooling, leveraging ambient conditions to reduce reliance on mechanical chillers. It can achieve up to 83.9 % energy savings compared to traditional CRAH units by utilising outdoor air directly or indirectly [61]. Colder, humid regions offer the highest efficiency, though indirect systems still deliver over 64 % savings in less ideal climates. For example, POLCOM Data Centre's modernisation incorporated free cooling below 6 °C, cutting annual energy use by 12 % [62]. These results highlight free cooling's adaptability across diverse climatic conditions, particularly temperate and arid regions.

Direct air-side economisers (DASE) use favourable outdoor conditions to reduce cooling loads, with optimised damper controls and dynamic supply adjustments cutting energy consumption by up to 46 % during winter weeks [63]. Free cooling remains viable even in hot regions (up to 40 °C) when humidity stays within acceptable thresholds. Managing supply air volume (SAV) in fan-wall systems further enhances thermal control, an SAV of 0.5 m^3/s per rack at a supply air temperature (SAT) of 27 °C reduces fan energy consumption by 34 % annually in mild climates [64]. Precise airflow design parameters, including a 60 % sieve porosity and 35 cm spacing between fan and sieve, can significantly reduce turbulence and cut airflow losses by 57.5 % [65]. These findings highlight the importance of airflow management within broader free cooling strategies. A comparative overview of major air-based cooling methods, including DASE, free cooling, and air-side economisation, is presented in Table 1.

Combining free air cooling with advanced technologies (i.e., evaporative cooling pads) can significantly boost data centre energy efficiency. In indirect free cooling (IFC) systems, evaporative pads enable effective operation even at ambient temperatures up to 42 °C and relative humidity below 20 %, maintaining a 20 °C temperature difference between dry cooler inlet and outlet water and achieving thermal efficiencies of 88 % to 90 % [72]. When airflow and water supply are optimised, these systems can reduce overall cooling energy consumption by 31 %. Leveraging favourable outdoor air conditions can push energy savings up to 50 %, supporting both economic and environmental performance [73].

Solar-driven solutions extend the applicability of free air cooling. A proposed solar chimney-based direct airside free cooling (SC-DAFC) system integrates natural ventilation with solar-induced buoyancy to maintain outlet air temperatures below 35 °C year-round in temperate regions [74]. Integrating a turbine within the chimney enables waste heat recovery and generates up to 15 kW of power, supporting both efficient cooling and sustainable energy use. Free air cooling also adapts to colder climates, as shown in subarctic regions like Luleå, Sweden. Optimising heat reuse, with exhaust temperatures of 40 °C to 50 °C, can raise energy reuse factors (ERF) from 0.50 to 0.66, while seasonal adjustments increase efficiency by 11–31 % [75], enabling both energy savings and secondary heat recovery.

Despite their advantages, free air cooling systems face challenges related to air quality, humidity control, and temperature variability, which can affect server reliability. However, advancements in environmental monitoring and control systems are mitigating these concerns. Additionally, hybrid configurations combining free air cooling with evaporative precooling and other techniques can enhance performance, achieving up to 40 % annual energy savings in temperate climates, depending on local climatic conditions and system configuration [76]. Overall, free air cooling is a sustainable and cost-effective approach. When integrated with technologies (i.e., evaporative pads, solar chimneys, and optimised airflow), it offers adaptable thermal management across diverse climates.

Indirect adiabatic cooling

Indirect adiabatic cooling (IAC) has emerged as an effective strategy for improving data centre energy efficiency, particularly in dry climates [77]. It works by spraying water into an airstream, where evaporation absorbs latent heat and lowers the air temperature under adiabatic conditions. This cooled air then passes via a crossflow heat exchanger to cool recirculated air from the data hall. The process avoids introducing humidity into the IT environment while leveraging ambient conditions. When combined with air-to-air heat exchangers, indirect evaporative cooling (IEC) systems can achieve up to 30 % energy savings over traditional mechanical cooling [78].

Integration of IAC systems into modern data centres significantly reduces the thermal load on mechanical chillers, lowering PUE values as low as 1.35 [79]. By pre-cooling intake air, IAC systems adapt well to varying climatic conditions. In cold, dry environments, characterised by ambient temperatures of 20 °C and humidity ratios of 6 g/kg, they can lower primary air temperature by up to 17 °C. Conversely, in hot and humid climates, with ambient temperatures of 40 °C and humidity ratios of 14 g/kg, they still achieve reductions of approximately 6 °C [80]. This versatility supports efficient thermal management across diverse climates and reduces reliance on energy-intensive mechanical cooling.

Adiabatic systems are particularly effective when integrated with free cooling, offering 30–40 % energy savings in warm, dry regions with relatively low water usage [81]. Hybrid configurations that combine IAC with vapour-compression systems enhance performance in hot, humid climates. For instance, cascading hybrid modes of indirect evaporative cooling have achieved up to 90 % cooling efficiency and 53 % energy savings compared to conventional systems [82]. These advancements underscore IAC's capability and effectiveness in enabling sustainable, energy-efficient data centre operations.

Advanced control strategies have elevated the efficiency and adaptability of IAC systems. Supervisory architectures, such as Newton-like phasor extremum seeking control (ESC), dynamically optimise air and water flow, reducing cooling energy consumption by up to 20 % compared to conventional control systems [83]. Pressurised water atomisers also offer precise temperature and humidity control, delivering 13 % to 32 % energy savings over traditional steam-based systems. These solutions are particularly valuable in sensitive settings (i.e., semiconductor facilities) where ultra-purified water ensures contamination-free operation [84].

Dew point indirect evaporative cooling (DPIEC) systems are advanced IAC technologies optimised for data centre applications. Counter-flow DPIEC designs can maintain supply air temperatures below 21.6 °C, with COPs exceeding 61.3 and cooling capacities over 15.5 kW/m^3 [85]. By extending free cooling periods and minimising reliance on mechanical refrigeration, they deliver substantial energy savings across various climates. These systems are often optimised using genetic algorithms and numerical simulations. As illustrated in Fig. 3, DPIEC uses a portion of the hot intake air to drive evaporative cooling through a counter-flow process, resulting in significantly lower supply air temperatures and improved energy efficiency.

Adiabatic cooling offers strong potential for enhancing energy efficiency and reducing operational costs in data centres. However, its performance depends on climate, airflow management, and water quality. Advancements in control systems, hybrid setups, and system design will further strengthen IAC's role as a sustainable cooling solution.

Direct adiabatic cooling

Direct adiabatic cooling (DAC), or direct evaporative cooling, is widely used in dry and arid climates for data centre thermal management [87]. It works by passing hot air through an evaporative medium, where water directly cools the air via evaporation before it enters the server room. This process reduces energy use compared to mechanical cooling and can lower air temperatures by 12–15 °C, depending on humidity and airflow, achieving energy savings of up to 30 % [88]. Advanced materials, such as cellulose-based pads, further enhance efficiency by reducing pressure drops and water consumption, enhancing DAC's sustainability and cost-effectiveness.

Table 1

Comparison of air cooling methods for data centres.

Cooling Method	Working Principle	Advantages	Limitations	Energy Efficiency (PUE, COP)	Climate Suitability	Implementation Cost & Payback Period	Case Studies & Experimental Data	Maintenance & Operational Challenges	Reference
Direct Free Cooling	Uses ambient air directly to cool IT equipment	– High energy efficiency – No refrigerants needed – Reduces chiller dependency	– Air contaminants require treatment – Humidity variations affect operation	PUE as low as 1.05 in optimal condition	Best for cool, dry climates	– Low-cost implementation – Savings depend on climate	Singapore testbed shows 38 % PUE reduction	Requires air filtration and periodic air quality checks	[66]
Indirect Free Cooling	Transfers heat via a heat exchanger without introducing outside air	– Avoids air contamination – Works well in a various climates	– Slightly lower efficiency due to heat exchange losses	Achieve 30–40 % improvement over conventional cooling	Works well in moderate temperature zones	– Moderate investment – Efficiency gains justify costs	Beijing study showed 10–20 % efficiency gains	Requires periodic heat exchanger maintenance	[67,68]
Air-Side Economisation	Uses external air to reduce mechanical cooling, mixing with recirculated air	– Reduces cooling costs by 40–50 % – Extends equipment lifespan	– Needs humidity – Condensation issues possible	Cooling energy savings up to 47 %	Best for mild climates	– Moderate investment – Significant operational savings	Europe and U.S. case studies show up to 67.2 % energy savings	Regular filter replacement required to prevent contamination	[69]
DASE	Introduces outside air with added filtration and humidity control	– Low operational cost – High energy efficiency	– Requires high-quality filters – Needs precise humidity monitoring	PUE reduced to ~ 1.1	Works best in temperate and dry climates	– Low investment – Significant long-term savings	U.S. study reports 47.5–67.2 % energy savings	Requires advanced filtration and humidity control	[69,70]
Evaporative Cooling Integration	Uses water evaporation to cool intake air	– High efficiency in hot, dry climates – Lowers mechanical cooling energy	– High water usage – Maintenance required for scaling and microbial growth	COP up to 52.5	Best for arid or semi-arid climates	– Higher upfront investment – Rapid payback in hot climates	China-based study reports COP improvement	Requires water treatment and regular system cleaning	[71]

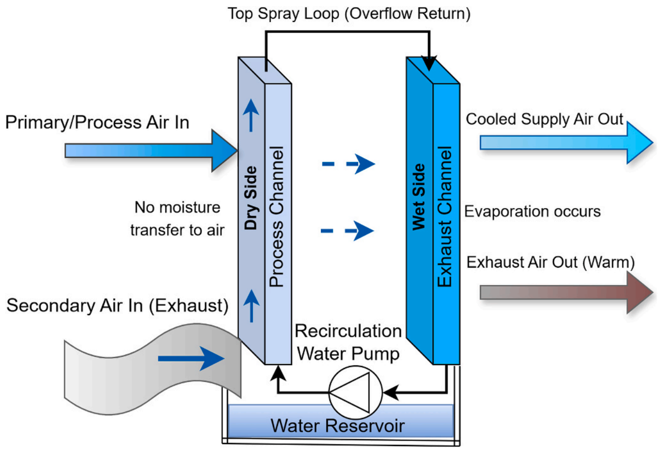

Fig. 3. Diagram of a Dew Point Indirect Evaporative Cooling (DPIEC) system, adapted from [86].

DAC performs best in hot, dry climates, where reduced humidity maximises cooling efficiency. Under optimal conditions, it can achieve over 85 % efficiency and reduce air temperature by up to 12 °C [89]. In humid environments, performance drops, but CFD models still show up to 8 °C temperature reduction with a 30 % increase in relative humidity, ensuring reliable thermal management [90]. Key challenges include scale formation, humidity control, and air quality. For example, excess humidity can impair electronics, necessitating precise control measures [91]. Innovations such as vertically split wet media and advanced evaporative pads improve durability and performance in demanding environments.

Cold-fog direct evaporative cooling (CDEC) is an advanced DAC principle that uses micro-droplet sprays to enhance heat transfer, particularly under low relative humidity conditions [92]. Unlike standard DAC, CDEC systems improve evaporation efficiency, lower supply air temperatures below 30 °C and meet ASHRAE data centre guidelines. Experimental results show 60 % to 98 % evaporation efficiency, depending on inlet air conditions [93]. Performance can further be optimised by adjusting spray flow rate, airspeed, and spray angle. Additionally, intermittent water supply strategies for evaporative pads help reduce water use without compromising cooling efficiency, balancing energy savings and operational sustainability [94].

Integrating direct evaporative cooling into data centres can significantly improve energy efficiency. For example, using porous evaporative media alongside air-conditioning systems achieves temperature drops of 5.8 °C to 15.9 °C, depending on pad thickness and ambient conditions [95]. This configuration improves COP by up to 44 % and reduces power consumption by approximately 20 %, making DAC highly effective in hot and dry climates. However, cooling efficiency is constrained by the wet-bulb temperature, which varies with humidity and temperature [96]. Reducing condenser air temperatures with evaporative pads further enhances system COP while maintaining low energy consumption.

DAC is less effective in humid climates (i.e., the U.S. Southeast) where high relative humidity reduces evaporation rates and limits cooling potential [97]. In such conditions, supply air temperatures may not meet the desired cooling requirements, and added moisture can exceed acceptable humidity levels for data centre operations [98]. DAC systems cannot cool air below the wet-bulb limit, which depends on initial air temperature and humidity levels. Mixing outdoor and indoor air raises contamination risks, potentially affecting equipment performance if not managed effectively. Additional challenges include scale formation on pads and water management, which must be addressed for long-term DAC sustainability.

Liquid-cooled data centres

Increasing power densities in HPC and AI environments are exposing the limits of traditional air cooling systems. As rack loads rise, air-based methods struggle with escalating heat fluxes, leading to inefficiencies and thermal instability. Advanced workloads can generate heat fluxes above 100 W/cm2, far beyond air's capacity, whose heat removal potential is only 37 % that of water [99]. This has accelerated the shift to liquid cooling. In 2020, average rack density reached 8.4 kW, with some exceeding 30 kW, compared to a legacy design built for 2 kW per rack [39]. This mismatch limits cooling capacities and scalability [100]. Without more efficient solutions, data centres face higher energy costs and overheating risks.

Liquid cooling offers a more effective alternative to air cooling by leveraging the higher thermal conductivity and heat capacity of water and dielectric fluids [102]. These systems manage high-power-density IT equipment more efficiently, especially in AI and HPC environments where air systems fall short [103]. Liquid cooling can be implemented in direct or indirect configurations: direct liquid cooling uses cold plates on CPUs/GPUs to extract heat at the source, while indirect liquid cooling (ILC) methods (i.e., immersion cooling) submerge components in thermally conductive fluids. Two-phase immersion cooling further enhances performance through phase-change heat transfer, improving heat dissipation for dense workloads.

This shift requires rethinking data centre design, as liquid cooling introduces challenges (i.e., high upfront costs, retrofitting complexities, and coolant leakage concerns). Integrating these systems into legacy facilities often involves significant structural modifications. As a result, adoption remains limited, with only a few hyperscale operators implementing liquid cooling at scale. Broader adoption will depend on overcoming costs, infrastructure, and standardisation barriers, critical as rack densities and thermal loads continue to climb.

Indirect liquid cooling

ILC addresses the thermal demands of high-performance data centres by transferring heat away from processors and memory using cold plates or other conductive interfaces integrated into liquid circuits. Unlike DLC, ILC utilises thermally conductive materials (i.e., silicone grease or welded interfaces) as an intermediate layer between the processor package and the coolant flow, avoiding direct coolant contact with the chip surface. This minimises leakage risks and makes it easier to retrofit existing facilities [104,105]. Water-cooled plates are the dominant ILC method in data centres with cabinet power densities between 20 and 50 kW, and over 90 % of liquid-cooled facilities use this approach [10,106]. ILC's compatibility with existing server architectures and lower structural demands make it a practical and scalable solution for AI and GPU-intensive workloads.

In addition to efficient heat dissipation, ILC reduces reliance on air-based systems, cutting energy use and enhancing component lifespan [107]. To further boost performance, He et al. developed a two-stage cooling system combining internal and external circuits, optimising flow rates and temperatures to improve dissipation [108]. Follow-up studies highlighted the importance of tailored coolant distribution strategies, which improved thermal stability and reduced energy consumption under variable workloads [109].

Sealed system innovations have further advanced ILC technology. Nicolas et al. developed a fully sealed internal loop, similar to a heat pipe, for localised heat management, paired with an external loop for heat rejection [110]. Their system effectively cooled two processors under varying loads using dynamic coolant flow control to maintain stable temperatures. While this design minimises leak contamination and thermal precision, scalability remains a challenge for hyperscale data centres.

Microchannel heat sinks are also critical to ILC systems, offering high heat flux in compact configurations. Yu and Cao highlighted their efficiency in fluid flow and heat transfer due to high surface area-to-volume ratios [111]. Zhou et al. introduced manifold microchannels to enhance flow distribution and thermal performance [112]. Additionally, liquid metal-based microchannels show promise for managing extreme heat fluxes in next-gen electronics [113]. However, broader adoption faces challenges related to material compatibility, corrosion risks, and cost, factors that must be addressed to meet the evolving cooling requirements of large-scale AI data centres.

Nanofluid-based cooling systems demonstrate significant potential for ILC applications, with studies reporting up to 47 % higher thermal conductivity than traditional coolants [114]. They also reduce pressure drops and enhance performance in microchannel heat sinks [115]. Additionally, hybrid approaches integrating nanofluids with conventional coolants have demonstrated improved heat transfer and lower pressure drops, making them suitable for high-power-density applications [116]. However, widespread adoption is limited by economic feasibility, long-term stability, and infrastructure [117]. Alkrush et al. further emphasised challenges (i.e., fluid degradation and microchannel clogging), which must be addressed through continued research before commercial-scale deployment is viable [118]. Embedded pin fins in microchannel heat sinks enhance turbulence and heat transfer while maintaining pressure drops manageable, ideal for high-power-density applications [119]. Advances in CFD modelling further support geometry optimisation, balancing thermal resistance and pumping power for peak performance [120].

Liquid metals have also shown promise for extreme thermal conditions. Zhang et al. demonstrated that liquid metal manifold microchannels can dissipate heat fluxes up to 1000 W/cm^2, maintaining chip temperatures below 351.7 K [121]. Other designs have achieved dissipation rates of as high as 1842 W/cm^2, far surpassing air cooling capabilities [122]. However, ILC systems still struggle to cool non-processor components like memory, power delivery, and storage devices, which also generate substantial heat [123]. Addressing these secondary thermal loads is essential to realising full-system liquid cooling solutions. In typical ILC systems, coolant distribution units (CDUs) deliver chilled liquid from external sources to internal loops that interface with servers, racks, and chips [124]. As AI workloads introduce rapid and uneven thermal fluctuations, dynamic control of CDU flow rates and thermal balancing mechanisms are required to ensure optimal performance.

While ILC offers energy savings, it requires tailored infrastructure. Meyer et al. noted that sealed enclosures and custom piping increase installation complexity and cost [125]. Tang et al. highlighted the need for higher pumping power due to increased viscosity and pressure drops, impacting hydraulic performance [126]. Environmental considerations also play a crucial role. Karimi et al. and Siddik et al. raised concerns about the indirect water consumption associated with liquid cooling, posing sustainability challenges, particularly in water-scarce regions [127,128]. Solutions include indirect evaporative systems that leverage free cooling potential in diverse environmental conditions while reducing energy demands [129,130]. Finally, Shi et al. proposed desiccant-based enhancements for hot-humid climates [131], while Bux et al. linked ILC adoption to reduced global warming potential in the digital era [132].

Despite its benefits, ILC adoption remains limited due to infrastructure modifications, high costs, and long-term reliability concerns. Overcoming these barriers is essential for widespread implementation in AI and hyperscale data centres.

Single-phase cooling

Single-phase cooling is an efficient and widely adopted solution for managing thermal loads in modern data centres. It offers a practical middle ground between air cooling and more complex two-phase cooling, combining energy efficiency, simplicity, and scalability with existing server architectures. Unlike two-phase cooling, single-phase systems circulate liquid without phase changes, maintaining a stable state throughout the process. This eliminates the need for vapour chambers or sealed environments, simplifying system design and scalability.

Recent advances in single-phase immersion cooling (SPIC) have highlighted its transformative potential for high-density computing. Wang et al. [133] showed that optimising operation strategies based on ambient temperature and chip power can reduce PUE by up to 16 %, especially under 600 W chip power. Ding et al. [134] explored enhancements such as micropump-induced flow turbulence and fan-assisted cooling techniques, improving efficiency by 21.39 %, though at the cost of increased system complexity and economic feasibility. Sun et al. [135] found that higher coolant flow rates reduce component temperatures but increase energy consumption and flow resistance. Similarly, Liu et al. [136] demonstrated a jet-assisted SPIC system that reduce peak CPU temperatures by 6.1 % and improved coolant temperature uniformity by 73.1 %, underscoring the potential of active cooling in performance gains.

The effectiveness of SPIC hinges on factors including pressure drop, volumetric flow rate, temperature uniformity, pumping power, and heat transfer coefficients. Optimised microchannel designs (i.e., HU-type and ZU-type manifolds) help balance these parameters, improving reliability and reducing energy demand [137]. Cheng et al. [138] also emphasised maintaining uniform temperatures and using high-conductivity materials (i.e., copper and aluminium) to improve thermal efficiency.

Advanced engineered coolants (i.e., Novec 7000) have been used to enhance heat transfer and reduce maintenance without introducing phase-change complexities [139]. However, 3 M has announced the phase-out of Novec fluids, including Novec 5110 and Novec 7000, due to environmental concerns related to PFAS (‘forever chemicals’), prompting alternative low-GWP dielectric coolants [140]. In response, vendors (i.e., EnviroTech Europe) have introduced PFAS-free alternatives like ProSolv®, offering similar thermal properties while ensuring compliance with low-GWP and evolving regulatory standards [141].

Coolant selection is critical across liquid cooling applications. Water-based fluids are commonly used in DLC systems where electrical isolation is not a concern. In contrast, dielectric fluids (i.e., hydrofluoroethers (HFEs), hydrofluoroolefins (HFOs), and fluoroketones (FKs)) are preferred for immersion or electrically sensitive environments due to their thermal stability and insulation properties [142]. For example, HFE-7100 and Novec 5110 have been widely used in single-phase immersion cooling, though Novec 5110 is now being phased out. Alternatives like Promosolv™ DR3 or Opteon™ are increasingly adopted. In two-phase systems, Opteon™ 2P50 and FK-based coolants offer low global warming potential and high heat transfer performance [143]. Spray cooling systems may also use dielectric oils or water-glycol blends, depending on the system design [144]. Table 2 summarises common fluid categories and their applications across major liquid cooling types.

Immersion cooling represents the most thermally efficient form of single-phase liquid cooling, achieving up to 100 % heat capture by fully submerging servers in dielectric fluids [146]. While its performance potential is unmatched, it remains an emerging technology due to substantial infrastructure demands. Applicability is currently limited to purpose-built data centres or specialised installations, as immersion cooling requires custom tanks, advanced fluid management, and structural modifications to support fluid containment and equipment weight.

Despite its advantages, single-phase cooling faces several challenges. Leakage risks, particularly with water-based coolants, remain in environments lacking robust containment. Flow maldistribution in microchannel designs can lead to thermal non-uniformity and hotspots, compromising reliability [137]. Moreover, single-phase cooling struggles to manage high-density heat loads exceeding 450 W/cm^2, limiting their scalability for hyperscale AI workloads [135]. Addressing these challenges requires improved containment, engineered fluids with higher thermal properties, and advanced microchannel geometries to enhance flow distribution and mitigate thermal non-uniformity. For instance, refined manifold designs and tailored flow paths can ensure more uniform temperature profiles across high-density components.

Table 2
Comparative evaluation of liquid cooling methods and typical coolants based on key performance metrics and fluid characteristics.

Cooling Type	Typical Fluid	Examples	Notes	Toxicity	Biodegradability	Maintenance	PUE Efficiency	Fluid Loss Risk	Cost
Direct-to-Chip (DLC)	Water-based or dielectric	Water, HFE, HFO	Dielectric fluids used for electrical isolation	★★★★☆	★★★★☆	★★★★☆	★★★☆☆	★★☆☆☆	★★☆☆☆
Immersion (1-phase)	Dielectric	HFE-7100, Novec 5110, Promosolv DR3	Novec 5110 is being phased out; consider replacements	★★★★★	★★★★☆	★★★★★	★★★☆☆	★★★☆☆	★★★☆☆
Immersion (2-phase)	Dielectric	HFOs (Opteon™ 2P50), FK fluids	Low-GWP alternatives gaining traction	★★★☆☆	★★★☆☆	★★★☆☆	★★★★★	★★★★☆	★★★★★
Spray Cooling	Dielectric	HFE-based oils, hydrocarbon mist	Common in motor windings and compact electronics	★★★★☆	★★★★☆	★★★★☆	★★★★☆	★★★☆☆	★★★☆☆

Ratings range from 1 star (★☆☆☆☆ = not very good) to 5 stars (★★★★★ = very good). Data adapted from [145].

Even with these challenges, single-phase cooling remains a promising pathway for high-performance electronic systems. Its energy efficiency, architectural compatibility, and scalability, particularly when enhanced through ongoing technological innovation, position it as a critical tool for addressing the growing thermal demands of AI-driven and data-intensive environments.

Two-phase cooling

Two-phase cooling is a highly efficient thermal management solution that utilises phase-change heat transfer, especially latent heat of vaporisation, to dissipate heat more effectively than single-phase systems [147]. This approach delivers thermal uniformity and high performance under extreme heat flux, making them suitable for demanding environments like data centres, aerospace, and microelectronics [148]. However, system complexity, material compatibility, and cost remain key barriers to widespread adoption. Ongoing research is focused on optimising configurations, working fluids, and system designs to address these limitations [149,150].

A key determinant of system performance is the choice of working fluid, which affects both efficiency and environmental sustainability. While perfluorocarbons (PFCs) offer thermal stability, their high Global Warming Potential (GWP) limits long-term viability [151,152]. Alternatives such as Hydrofluoroolefins (HFOs) like TMC-49 and fluoroketones (FKs) like Novec 649 provide low to ultra-low GWP ($\sim$1), nonflammability, and favourable thermal properties [153,154]. These fluids are increasingly employed in two-phase immersion systems, though they require stringent filtration and fluid hygiene to ensure long-term stability.

The choice of working fluids in two-phase immersion cooling involves balancing thermal performance, environmental impact, and system design. PFCs, though thermally stable, are largely confined to legacy or sealed systems due to their high GWP. HFEs offer a reliable balance of thermal performance and moderate sustainability. Modern alternatives such as HFOs (e.g., Opteon™ 2P50) and PFAS-free fluoroketones provide ultra-low GWP and strong dielectric properties, making them well-suited for ESG-aligned, high-performance applications such as AI and HPC [155].

Several studies have investigated two-phase cooling mechanisms and optimisation strategies. Tong et al. [156] evaluated two-phase thermosyphon loops (TPTLs) using CO_2, R134a, and R410A for data center applications. The findings demonstrated that CO_2 systems exhibited lower thermal resistance and higher heat transfer coefficients, with optimal filling ratios of 45 % for CO_2 and 35 % for R134a/R410A. Unlike the flow instabilities observed with other refrigerants, CO_2 exhibited oscillatory behaviour that enhanced thermal stability. However, the systems' sensitivity to filling ratios and design complexity limits their scalability in AI-driven data centres requiring precise thermal control.

Wu et al. [157] analysed two-phase LIC for extreme heat loads, showing effective temperature regulation under chip power exceeding 600 W. Stable operation depends on boiling and condensation dynamics, but dry-out risks at high heat fluxes and the need for precise coolant circulation complicate large-scale implementation. Despite these limitations, LIC remains a promising solution for high-density computing environments.

Pool boiling, a subset of two-phase immersion cooling, is particularly effective in managing heat in high-power-density electronics. It enables continuous rewetting of heat sink surfaces through phase change, maintaining low, uniform temperatures with minimal moving parts [158]. However, high wettability in dielectric liquids can inhibit vapour trapping, increasing incipient superheats. To overcome this, surface engineering techniques, including micro- and nano-structured features, have been employed to enhance vapour pathways, increase nucleation, and enhance liquid replenishment. Porous and textured surfaces further improve boiling efficiency and thermal sustainability [159].

Complementing these efforts, phase change materials (PCMs) are being explored as thermal buffers in data centres [160]. PCMs absorb and release heat during phase transitions, enabling improved temperature control and energy savings. While challenges remain, particularly in material selection, scalability, and economic feasibility, ongoing research into PCM optimisation and real-world validation could unlock new pathways for passive thermal management, contributing to carbon–neutral, energy-efficient data centre operations.

A related study investigated nanowire spacing using FC-72 as a dielectric fluid [161]. The findings indicated that increased spacing led to larger cavities, enhancing liquid flow and reducing incipience superheat while increasing critical heat flux (CHF) and heat transfer coefficient (HTC). Copper nanowires exhibited slightly better performance due to higher surface roughness, although silver-coated surfaces yielded comparable results. These findings highlight the potential of nanostructured coatings in optimising boiling performance for immersion-cooled electronics.

Mechanically driven two-phase cooling loops (MTCLs) offer energy-efficient operation, particularly when incorporated with free cooling. Gong et al. [162] reported annual energy savings of up to 25.78 % while maintaining consistent performance under high heat flux. However, slow startup times and efficiency fluctuations remain challenges for large-scale AI applications.

Economic feasibility is a key to adopting two-phase cooling. Kanbur et al. [163] found that two-phase immersion systems achieve higher COP and lower energy consumption than single-phase systems. Yet high initial costs, due to dielectric fluids and complex system architectures, limit widespread implementation. In addition, the maintenance complexity and precision required make it difficult for legacy data centres to transition from traditional air or single-phase cooling solutions.

Recent design advancements have further improved two-phase cooling performance. Sun et al. [164] demonstrated that integrating baffles and novel coolants like Novec 7000 enhanced heat dissipation and temperature uniformity. Zhou et al. [165] showed that capillary wicking structures reduced thermal resistance to 0.0454 °C/W at 600 W, underscoring the importance of precise system configuration. However, reliance on specialised materials and infrastructure increases implementation costs, limiting adoption in hyperscale facilities.

Immersion cooling has been shown to reduce energy consumption by up to 50 % while requiring two-thirds less physical space compared to traditional air cooling, enabling higher rack power densities [166]. Studies reveal two-phase immersion and cold plate systems outperform air cooling by achieving up to 51 % higher CPU clock rates and lower thermal resistance [167]. Advanced structures such as microchannels and conical enhancements further improve efficiency [168]. Material improvements are also advancing cooling performance. Polymer-based composites, enhanced with bio-based or synthetic additives, have improved thermal and mechanical properties for electronics and energy storage systems [169]. These developments align with broader sustainability goals.

Overall, two-phase cooling represents a transformative approach for high-density environments, combining superior heat dissipation, adaptability, and energy efficiency. To enable broader adoption of AI and hyperscale data centres, the remaining challenges in system complexity, cost, and control precision must be addressed.

Heat-pipe cooling

Heat-pipe technology is increasingly recognised as an efficient thermal management solution for data centres, leveraging phase-change heat transfer to dissipate heat rapidly with minimal thermal resistance. These systems operate by evaporating a working fluid at the heat source and condensing it at the sink, using capillary or gravitational action to return the fluid. Their passive nature and compact design make them attractive for AI accelerators, GPU clusters, and HPC environments, where thermal stability is crucial. However, scalability, material compatibility, and operational complexity remain key barriers in hyperscale deployments.

Recent studies have explored various configurations for the performance enhancement of heat-pipe technology. Wang et al. [170] demonstrated that loop heat pipe heat sink (LHPHS) reduced thermal resistance to 0.044°C/W with stable temperature differentials of 0.265 °C, but required higher coolant flow rates, increasing energy consumption. She et al. [171] introduced a hybrid system combining separate heat pipes (SHP) with adsorption refrigeration (AR) to recover waste heat. It achieved PUE values between 1.082 and 1.095, though performance was climate-dependent, and system complexity limited scalability in tropical and hyperscale environments. Sun et al. [172] proposed a pump-driven heat pipe/vapour compression (PHP/VC) hybrid system for rack-level cooling, achieving an impressive annual average COP of 7.81. While highly energy-efficient, the system's reliance on precise pump and fan control adds operational complexity, limiting practicality for hyperscale AI deployments.

Structural advancements have also enhanced heat-pipe performance. Zhang et al. [173] studied flat loop heat pipe (FLHP) configurations and found that mid-channel wick designs stabilised vapour-liquid circulation and reduced temperature fluctuations. However, their use of R141b, a high-GWP fluid, raised environmental concerns, emphasising the need for sustainable alternatives. He et al. [174] applied genetic algorithms to optimise integrated heat pipe (IHP) systems combining natural cooling, vapour compression, and integrated modes. These designs improved energy efficiency by 2–3 times, with 16.18 % to 20.08 % energy savings under varying climates. Still, their complexity and reliance on accurate ambient temperature predictions limit their real-world applicability.

For chip-level cooling, Xue et al. [175] explored loop heat pipes (LHPs) and achieved PUE values of 1.112 and 1.076, operating effectively at higher inlet water temperatures and reducing chiller dependence. However, peak summer conditions and the need for precise thermal management present deployment challenges. Similarly, Weng et al. [176] valuated micro-channel flat loop heat pipes (MCFLHP), reporting a peak heat recovery efficiency of 84.06 %. Their modular and compact design offers scalability for large data centres, but the performance was highly sensitive to coolant flow rates and external heat recovery setups, requiring adaptable and robust system designs for reliable operation.

Hybrid heat pipe systems have shown promising energy efficiency results. Du et al. [177] integrated hybrid water-cooled loop heat pipes with CRAC systems, achieving 34.9 %-52.4 % annual energy savings, though high initial costs and reliance on natural cooling restricted their geographical feasibility. Tang et al. [178] enhanced flat loop heat pipes with Tesla valves, improving thermal performance but increasing dependence on water infrastructure, limiting applicability in water-scarce regions. Xiang et al. [179] advanced micro-channel flat loop heat pipes (MC-FLHP) designs, achieving thermal recovery efficiencies up to 84.52 % by improving evaporator and condenser components. However, high thermal load management and low inlet water temperature requirements remain challenges. Li et al. [180] examined aluminium flat plate pulsating heat pipe modules for chip-level cooling, effectively managing heat loads up to 250 W and outperforming conventional air cooling. Their findings emphasised the importance of material selection and design improvements to support scalable, high-performance cooling in AI-driven data centres.

Despite recent advancements, heat-pipe technology faces inherent limitations. Its passive nature, while reliable, restricts its ability to handle extreme heat loads compared to active systems. Additionally, its performance is highly dependent on design parameters such as coolant selection, pipe geometry, and operating conditions, necessitating careful customisation that complicates large-scale deployments. Maintaining consistent performance across varying environments, especially under fluctuating heat loads, remains a major challenge. Furthermore, environmental concerns related to refrigerants (i.e., R141b) underline the need for sustainable working fluids that offer high thermal performance without ecological trade-offs.

Direct-to-chip liquid cooling

DLC has emerged as a leading technology for thermal management in HPC and data centre environments. By applying dielectric fluids directly to heat-generating components, DLC enables precise and localised cooling. Unlike air-based techniques that rely on large-scale airflow and heat sinks, DLC allows direct thermal contact with processors and other critical electronics, significantly improving thermal efficiency and system reliability. Technically, DLC systems achieve a heat capture efficiency in the range of 70–75 %, allowing them to handle significantly higher thermal loads than conventional cooling systems [181].

A key advantage of DLC is its ability to reduce cooling energy consumption while maintaining uniform temperature profiles. By eliminating the need for conventional chillers and reducing airflow requirements, it enables substantial energy savings [182]. Phase-change mechanisms in DLC systems stabilise heat flux, improving thermal performance and operational reliability. When coupled with PCMs and three-dimensional oscillating heat pipes (3D-OHPs), DLC has been shown to reduce surface temperatures by approximately 35°C compared to air cooling [183]. Eliminating airflow dependency also allows for greater rack configuration flexibility, reduced mechanical stress, and lower noise levels.

Beyond energy efficiency, DLC provides significant performance advantages in high-density environments where air cooling struggles to meet thermal demands [184]. Its localised, uniform heat removal improves thermal stability and increases component reliability. Unlike some other liquid systems, DLC eliminates the need for sealed enclosures and piping, improving system adaptability. Additionally, phase-change effects within DLC systems further stabilise temperature profiles, supporting consistent operation under high thermal loads.

Despite its advantages, DLC adoption faces several challenges. Dielectric fluids provide excellent electrical insulation but generally have lower heat transfer efficiency than water-based coolants, creating a trade-off between electrical safety and thermal performance.

Maintenance is also complex as DLC systems require hermetically sealed designs to prevent fluid loss, air infiltration, and humidity-related degradation. These design constraints complicate servicing, particularly in hot-swapping scenarios, where minimising downtime is critical in high-availability data centre environments.

Economic barriers also hinder DLC implementation. High upfront costs, including specialised pumps, cooling circuits, and auxiliary heat exchangers, can deter operators despite DLC's long-term energy savings. Additionally, supporting infrastructure like secondary cooling loops and pumps introduces extra energy demands. Although DLC can be integrated into existing data centres, it often requires layout modifications or facility-level retrofits to accommodate CDUs and coolant routing, limiting its plug-and-play potential [182].

Scalability remains a key concern for DLC, particularly in hyperscale data centres, where integration often requires extensive architectural modifications, increasing implementation costs and deployment times. Safety risks associated with fluid leakage and high pressures in two-phase systems further complicate adoptions [185]. However, emerging hybrid strategies, such as combining DLC with heat pipes or PCMs, offer promise for improving thermal efficiency while mitigating these limitations [186].

Despite these constraints, adoption of DLC systems is steadily growing, driven by their suitability for AI clusters, dense compute racks, and energy-efficient facilities. DLC is gaining traction as a thermal management solution for HPC, AI workloads, and energy-efficient data centres. While economic and operational challenges remain, advancements in fluid formulations, system architectures, and hybrid integration strategies are driving their adoption in next-generation cooling technologies.

Cold plate liquid cooling

Cold plate cooling is a highly efficient form of DLC, involving metal plates attached directly to heat-generating components (i.e., CPUs and GPUs) [187]. These plates increase heat sink surface area and use a closed-loop liquid system to remove heat, ensuring superior thermal stability and dissipation than air-cooled heat sinks [188].

Cold plate cooling offers substantial energy savings. By reducing airflow and eliminating chillers, they can cut cooling energy consumption by over 90 %, and computational energy use by up to 14 % compared to air-cooled systems [189]. Full cold plate deployment has achieved 18.1 % facility power savings and a 10.2 % total power reduction compared to 100 % air-cooling solutions [190]. A defining feature of cold plate systems is their high thermal efficiency, particularly in direct-contact applications. Closed-loop DLC systems can achieve a temperature differential of just 15 °C between CPU and coolant, enabling effective heat transfer and stable performance [191]. This is especially critical in AI, HPC, and high-density data centres, where thermal control directly impacts system reliability and longevity.

Cold plate cooling has been successfully deployed in next-generation liquid-cooled supercomputers. For instance, Hewlett Packard Enterprise (HPE) developed a 100 % fanless DLC architecture utilising cold plates and fluid loops to extract heat directly from all components, significantly reducing cooling energy use and operational costs [192]. These deployments demonstrate DLC's potential to support high-density, energy-efficient data centres while maintaining system reliability [193]. Such advancements underscore the growing role of cold plate technology in enhancing computational efficiency and reducing data centre energy footprints.

Despite its advantages, cold plate cooling faces several challenges limiting broader adoption. The need for specialised components, including sealed piping, liquid-cooled heat exchangers, and leak detection systems, increases installation and maintenance costs. Additionally, cold plates require secondary heat rejection systems (e.g., liquid-to-liquid or liquid-to-air heat exchangers), increasing operational complexity and energy demands.

Scalability is particularly difficult in retrofit scenarios, where integration often requires substantial modifications to existing infrastructure [194]. Adapting air-optimised spaces to support cold plate systems may involve new piping networks, re-engineered server racks, and environmental monitoring. Additionally, cold plates typically handle only 60–70 % of total heat dissipation, necessitating supplemental chilled air systems to manage residual thermal loads [195]. The need for hybrid cooling approaches further increases system complexity and total cost. The key operational limitations of cold plate cooling, including cooling capacity constraints, efficiency challenges, design complexity, environmental risks, and maintenance requirements, are summarised in Table 3.

Hybrid strategies that combine cold plates with technologies (i.e., heat pipes, PCMs, and thermoelectric elements) are being explored to enhance efficiency and overcome existing limitations. These integrated approaches optimise performance and reduce energy consumption, making cold plate cooling increasingly attractive for HPC, hyperscale data centres, and AI-driven workloads.

Spray cooling

Spray cooling utilises atomised droplets to directly cool hot surfaces, rapidly dissipating heat through evaporation and convection mechanisms. As a form of two-phase flow boiling, spray cooling offers high heat transfer efficiency by leveraging phase change directly on the surface of electronic components, enhancing thermal resistance and improving heat convection [200]. This makes it particularly effective for high-power density servers in data centres and HPC environments.

A key advantage of spray cooling is its ability to maintain temperature uniformity across electronic components, ensuring reliable performance under intense thermal loads [201]. For instance, spray-cooled

Table 3

Operational limitations of cold plate liquid cooling.

Limitation Category	Description	Impact	Source
Cooling Capacity	Maximum rack density of 50–60 kW	Unsuitable for next-gen high-density AI applications	[195]
	Handles only 60–70 % of server heat	Requires supplemental air cooling	
Efficiency	PUE of around 1.15	Less efficient than advanced alternatives like immersion cooling	[196]
	Need for additional air conditioning	Increases energy consumption and operational costs	
Design and Integration	Difficulty fitting cold plates to all components	Limits effectiveness in high-density servers	[197]
	Increased complexity at scale	Complicates installation, maintenance, and troubleshooting	
Environmental and Reliability	IT assets exposed to environmental factors	Can accelerate component deterioration	[198]
	Risk of coolant leaks	Potential for catastrophic damage to electrical components	
Maintenance	Requires frequent visual inspections	Increases operational overhead	[199]
	Coolant quality management	Necessitates regular monitoring and maintenance	
	Flow and pressure monitoring	Requires ongoing adjustments for optimal performance	

rack systems using near-ambient coolant have outperformed conventional air cooling, eliminating the need for mechanical chillers. Additionally, applying spray evaporative cooling to air-cooled chiller condensers has increased COP by 4–8 % and reduced electricity consumption by 2.37–13.53 % [202].

Despite its advantages, integrating spray cooling into existing data centre infrastructure presents challenges. Effective deployment requires careful consideration of coolant flow dynamics and heat transfer characteristics. In parallel two-phase loops, nonuniform heating can lead to flow maldistribution and local dry-out in cold plates, even at inadequate overall flow rates [203]. While flow restrictors can mitigate this issue, they increase system complexity and pressure drops, complicating large-scale implementation.

System stability under varying thermal loads is another significant challenge for spray cooling. These systems must dynamically adapt to changing server workloads to avoid thermal throttling and ensure consistent performance [204]. Ensuring stable pressure and temperature is also critical, as fluctuations can disrupt boiling efficiency [201]. Moreover, coolant leakage remains a safety concern which requires reliable connectors and robust system design to protect electronics and prevent potential data loss [205].

Spray cooling's effectiveness depends highly on nozzle design, spray flow rate, and nozzle-to-surface distance, which influence droplet size, impact velocity, and liquid film behaviour, all critical to optimising heat transfer [206]. Advanced nozzle designs, including pressure-swirl and straight-tube configurations, enhance atomisation and cooling capacity, particularly in cryogenic and high-power applications. Adjusting parameters, such as spray angle, can also mitigate stagnation zones, improving overall heat dissipation efficiency.

Nanofluids are a promising innovation in spray cooling, with studies demonstrating that incorporating graphene oxide (GO) or aluminium oxide (Al_2O_3) enhances thermal conductivity and surface wettability, boosting heat transfer coefficients [207,208]. However, performance is concentration-dependent, excessive nanoparticles increase viscosity, reducing flowability and system efficiency. Achieving optimal results requires careful tuning of nanoparticle properties and spray [208].

A study by [209] highlights the importance of atomisation characteristics in spray cooling. Smaller droplets and higher velocities improve heat dissipation but can reduce uniformity due to velocity variations, suggesting the need for advanced nozzle designs that maintain consistent spray behaviour. Additionally, nano-alumina additives can enhance convective heat transfer at optimal concentrations, although excessive use may increase thermal resistance due to particle stacking.

Hybrid nanofluids are being explored as next-generation thermal fluids for spray cooling, capable of managing heat fluxes from 100 to 1000 W/cm^2 [207]. They offer improved performance over traditional coolants but require careful balancing of nanoparticle concentration, fluid viscosity, and spray parameters to optimise cooling without introducing excessive thermal resistance.

Spray cooling systems have demonstrated significant potential in sustainable thermal management. For example, two-phase systems using near-ambient coolant temperature ($\sim$30 °C) have outperformed air cooling, eliminating the need for chillers and demonstrating durability under high-temperature conditions [201]. Full-scale spray-cooled rack systems confirm their scalability, although ambient temperature variations can impact cooling efficiency and pressure stability [210]. Despite these advances, spray cooling is still not a mainstream solution in data centre deployments due to technical and industrial barriers. Key barriers preventing the widespread adoption of spray cooling:

1. Leakage: Fluid leaks can cause electrical shorts or hardware damage.
2. Complexity: Managing spray flow, droplet uniformity, and return paths for liquid is technically demanding.
3. Reliability: The Risk of nozzle clogging or uneven cooling reduces system dependability.
4. Material Compatibility: Cooling liquid integration with sensitive electronics remains challenging.
5. Standardisation: Lack of standard industrial frameworks hinders scalable deployment.

A study on secondary cooling in continuous slab casting offers transferable insights for data centres' spray cooling. It found that smaller droplets improve cooling uniformity, while larger droplets enhance cooling intensity but reduce coverage [211]. Nozzle spacing and droplet velocity were also identified as critical for optimising efficiency, highlighting the need for precise control over nozzle configuration and spray dynamics.

Innovative configurations, such as cold-mist direct evaporative cooling (CDEC), provide additional opportunities for data centres' energy efficiency. Optimising spray angles and airspeeds has been shown to reduce air temperatures and improve humidity control, achieving 14–41 % energy savings compared to modular air-cooled heat pumps [212]. However, humidity control under high temperatures and adaptability to varying climate variability remain key challenges.

Computational models have been employed to optimise droplet size and flow parameters in spray cooling. Droplets under < 20 μm improve evaporation rates and reduce water consumption, but practical implementation requires consideration of droplet collisions and secondary breakup effects [213]. Similarly, the integration of falling film evaporators with modified tube surfaces and baffle plates has also boosted heat transfer, achieving 3.4–3.5 times higher coefficients than traditional setups [214].

While spray cooling offers a transformative approach to thermal management for data centres and other high-power-density applications, challenges remain. Effective deployment requires precise control of droplet size, impact velocity, and nozzle positioning to prevent uneven cooling or excessive liquid accumulation. Scalability, droplet stability, and coolant recovery systems must be addressed to ensure efficient and safe operation at scale.

Jet impingement

Jet impingement cooling is increasingly recognised for its precision and efficiency in thermal management, particularly in compact data centres and high-power electronic systems. By targeting high-velocity fluid jets onto localised hotspots, it ensures uniform heat dissipation, making it ideal for AI accelerators, GPUs, and HPC components. Advances in nozzle geometry, flow uniformity, and hybrid integration have significantly expanded its real-world deployments.

System efficiency depends on parameters such as jet diameter, nozzle-to-surface spacing, flow rate, and Reynolds number. Experimental and computational analyses show that optimising these parameters improves heat transfer and reduces thermal resistance. For instance, smaller nozzle gaps enhance convective but may slightly increase pressure drops [215]. Arrays of confined jets have been demonstrated to reduce junction temperatures and improve temperature uniformity in dense semiconductor modules.

Recent research has explored multi-jet systems using advanced coolants. HFE-347 has shown superior performance over white mineral oil, achieving higher heat transfer coefficients and thermal uniformity [216]. Hybrid systems combining jet impingement with PCMs or microchannel cooling have further advanced the technology to manage heat fluxes exceeding 2100 W/cm^2 [217,218]. These systems provide both localised cooling and system-wide thermal stability, key for next-generation semiconductor devices.

Recent advances in jet impingement cooling have introduced several high-potential technologies for advanced data centres. Liquid synthetic jet devices (LSJDs), driven by piezoelectric actuators, generate high-velocity jets, localised cooling with heat transfer coefficients up to 1.53 W/cm^2·K [219]. Optimised configurations, such as an orifice-to-heater distance ratio of 4, leverage vortical penetration into cross-flows, offering a 12-fold performance gain over conventional liquid cooling [219]. LSJDs are particularly suited for immersion-cooled, silent, and energy-efficient data centres.

Rotary jet impingement is another technique that utilises rotating pipes with embedded nozzles to ensure uniform cooling over cylindrical surfaces. Increasing Reynolds numbers and optimising rotational speeds improve performance, with Nusselt number peaks observed in stagnation zones, emphasising the importance of nozzle geometry and jet-to-surface distances [220,221].

Synthetic jet systems, which use oscillatory motion rather than continuous flow, offer compact cooling for localised hotspots. Studies show that maximum heat transfer occurs at jet-to-surface distances of about five times the jet diameter, through warm air recirculation can reduce cooling uniformity at small spacings [222]. Optimising actuator configurations and jet velocity helps mitigate these issues.

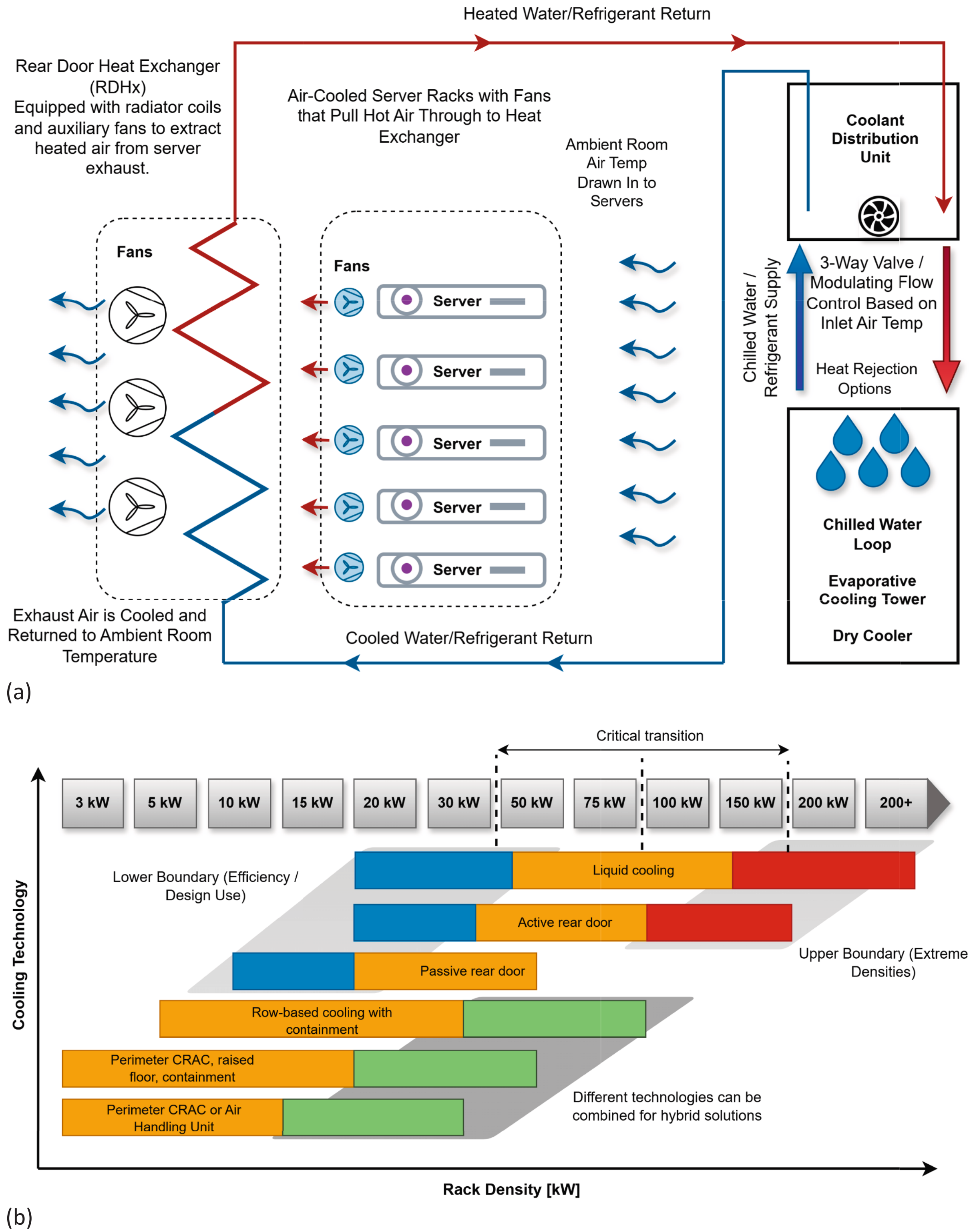

Fig. 4. (a) Schematic of a RDHx cooling system illustrating rack-level liquid cooling integration. Adapted from [229]. (b) Applicable cooling technologies mapped against rack power density, highlighting deployment suitability across air, hybrid, and liquid systems. Reproduced from [101].

Despite the advancements, flow uniformity across multi-jets remains a challenge. Designs inspired by the Coanda effect show promise in redirecting jet flow for improved velocity distribution, though issues like pressure drops and maldistribution persist [223]. Studies on Z-shaped inlet/outlet geometries have demonstrated the ability to reduce thermal hotspots and stabilise cooling in high-density applications [224].

Outlet configuration plays a critical role in optimising jet impingement systems. Research indicates that reducing outlet sizes by 30 % can decrease spatial temperature differences by up to 31 % and improve hotspot heat transfer coefficients by 20.5 % [225]. Circular top outlets offer a balance between temperature uniformity and pressure drop, whereas side-outlet configurations improve temperature variations but result in higher pressure losses due to longer fluid paths.

Advancements in internal flow architectures further enhance performance. Dual- and triple-chamber cold plate designs with jets, xtended extraction pathways, enhance turbulence and heat transfer rates while maintaining manageable pressure drops [226]. Curved walls and extended jet paths have been shown to reduce thermal resistance and improve temperature uniformity, making them effective for modern data centre cooling.

Finally, despite significant advances in nozzle design, synthetic jets, rotary systems, and hybrid configurations, challenges remain, particularly velocity uniformity, maldistribution, and pressure optimisation. Continued refinement of multi-jet configurations, internal geometries, and outlet strategies will be key to realising jet impingement cooling's full potential for compact, high-power systems.

Rear door heat exchangers (RDHx)

Rear Door Heat Exchangers (RDHx) are efficient, rack-level cooling solutions designed to manage the increasing thermal demands of high-density IT equipment. Mounted at the rear of server racks, RDHx systems use air-to-liquid heat exchangers to capture and dissipate heat as it exits the servers [227]. As shown in Fig. 4(a), hot exhaust air passes through a liquid-cooled coil, transferring heat to a circulating coolant loop routed through a CDU and expelled via chillers, dry coolers, or evaporative towers. RDHx offers a modular, lower-barrier alternative to full DLC systems and is well-suited for brownfield upgrades due to its compatibility with traditional data centre layouts. Technically, RDHx systems can achieve up to 100 % heat capture efficiency by directly removing heat from server exhaust air. This high efficiency, combined with their retrofit-friendly nature, has led to increasing adoption in both colocation and hyperscale facilities [228].

This RDHx schematic highlights how rack-level thermal isolation and higher return water temperatures improve energy efficiency and facilitate integration with hybrid systems. By localising heat removal, RDHx minimises thermal losses, reduces reliance on traditional air conditioning, and eliminates the need for hot and cold aisle separation. Shortening air distribution paths achieves higher cooling efficiency even at supply-air temperatures of 24 °C [230]. CFD analyses demonstrate that RDHx can enable chilled water temperatures up to 14 °C, improving secondary cooling performance. For example, in a 30 MW reference data centre, RDHx systems achieved a PUE cooling range of 1.25–1.33, with energy use distributed across chillers (46 %), pumps (33 %), and RDHx fans (21 %), highlighting the system's energy-efficient design.

Experimental studies confirm that RDHx systems significantly improve cooling performance and energy savings. Liquid-cooled cabinets using RDHx have demonstrated a 66 % increase in cooling efficiency and lower operating temperatures compared to conventional air-cooled counterparts [231]. Similarly, annual energy savings of over 30 % have been reported, highlighting RDHx's stability for hybrid and free cooling strategies [232]. By transferring server exhaust heat directly to a cooling fluid, RDHx reduces the mixing of hot and cold air, a key limitation of traditional CRAC systems. For example, rack-level heat pipe integration has reduced entransy dissipation by 25.4 % compared to CRAC systems, supporting RDHx adoption in high-density environments [233].

In hybrid cooling configurations, RDHx combines air and liquid cooling by using finned heat exchangers and multiple fans to enhance heat removal. The warmed water exiting the heat exchangers can be reused for direct-to-chip liquid cooling, improving overall system efficiency [234]. Integrating heat pipe technology further optimises performance via phase-change heat transfer, enabling stable operation at higher ambient temperatures of up to 15 °C. Studies show that optimising condenser wind speed and refrigerant pump frequency improves energy efficiency under varying climatic conditions, making RDHx an adaptable and reliable solution for diverse data centre environments [235].

RDHx systems contribute not only to efficiency but also to sustainability. In the Frontier HPC-DC, RDHx systems are integrated with high-temperature heat pumps (HTHPs) to recover waste heat, reducing reliance on chillers and supporting carbon reduction goals [236]. By addressing thermal loads at the rack level, RDHx helps maintain optimal conditions while lowering operational costs and emissions. Recent advancements (i.e., optimising coolant flow rates and heat exchanger designs) have further improved RDHx performance and reliability under varying server workloads [231].

Despite their advantages, RDHx systems face scalability and operational complexity challenges. Their transient response and time constant are critical for thermal control optimisation during load variations [237]. Configuring gravity- or pump-driven refrigerant flows requires balancing energy efficiency with installation compactness [238]. While high upfront costs, particularly in hybrid configurations, may limit adoption in smaller data centres, their versatility and ability to be added to existing infrastructure without major modifications make them a compelling option for modern high-density environments. Compared to traditional CRAC and row-based containment systems, RDHx offers a scalable midpoint between air and direct-to-chip liquid cooling. As shown in Fig. 4(b), its operational envelope spans moderate to high rack power densities, overlapping with both passive rear doors and liquid-based solutions, making it a versatile choice in hybrid deployments.

Commercial readiness and industry deployment

This section provides a strategic overview of the commercial readiness, technological maturity, and business relevance of advanced cooling technologies in the data centre industry. While earlier sections focused on technical performance and innovation, the emphasis here is on real-world deployment, highlighting which solutions are in use, who is adopting them, and under what conditions they succeed. It also explores market ecosystems, vendor landscapes, and investment potential, offering guidance for both technical planners and business leaders.

Air-based cooling systems: current status and transitional role

Air-based cooling continues to dominate thermal management in the data centre industry, particularly across the enterprise, colocation, and small-to-mid-scale facilities. These systems, which include CRAC and CRAH units, collectively account for approximately 88 % of total data centre cooling market revenues [239]. Their commercial success stems from their simplicity, cost-effectiveness, and compatibility with legacy infrastructure, making them a preferred choice for standard-density applications where rack loads typically remain below 12 to 15 kW [240].

CRAC systems are widely deployed in lower-capacity facilities due to their minimal installation requirements and direct air-based design. CRAH units, on the other hand, offer enhanced energy efficiency through the use of chilled water systems and are better suited to larger-scale deployments. These units are often integrated with aisle containment and economiser strategies to extend their operational efficiency. Leading operators such as Equinix and Telehouse have implemented such systems as part of broader airflow optimisation and sustainability strategies [241].

In regions with favourable ambient conditions, particularly cold or arid climates, air-based cooling has evolved to incorporate advanced free cooling and evaporative methods. These systems pre-condition or substitute mechanical cooling with filtered outdoor air or water-assisted cooling, reducing reliance on compressors and chillers. Hyperscale cloud providers, including Microsoft, Google, and Amazon, have adopted such climate-responsive cooling strategies to minimise energy and water use across their global campuses [242]. In Northern Europe, for instance, Microsoft reports saving over 125 million litres of water annually through the use of indirect free air systems [243]. Microsoft also aims to cut water usage in evaporative-cooled data centres by 95 % by 2024 by favouring adiabatic cooling in cooler regions, though scaling such systems still poses water scarcity challenges [244]. Fig. 5 illustrates how regional climate significantly influences the viability and efficiency of air-based cooling systems, particularly those relying on direct or indirect use of ambient air. While these approaches deliver substantial operational savings and environmental benefits, their effectiveness remains geographically constrained due to sensitivities to humidity, air quality, and seasonal variation.

Despite continued improvements, air-based systems face growing performance limitations in the face of modern compute demands. Fan energy consumption accounts for a non-negligible proportion of facility power use, often as high as 15 %, and the thermal stability of air cooling deteriorates rapidly as rack densities increase beyond 20 kW [245]. In emerging workloads such as artificial intelligence training, high-performance computing, and GPU-dense server configurations, heat fluxes are frequently approaching thresholds that conventional air systems cannot support. As a result, industry momentum is gradually shifting toward higher-efficiency alternatives capable of managing elevated thermal densities.

Nonetheless, air cooling remains commercially viable and widely deployed. According to a 2024 industry survey more than 60 % of data centre operators plan to continue relying primarily on air cooling through at least 2026 [246]. This persistence reflects its enduring advantages, low capital expenditure, ease of maintenance, and integration flexibility. However, there is an increasing trend toward hybrid cooling architectures, particularly in newly built or retrofitted facilities seeking to accommodate future thermal scalability.

Looking ahead, the role of air cooling is expected to transition from a standalone backbone to a complementary component within hybrid environments. While no longer sufficient for the most thermally intensive applications, air-based systems will continue to play a strategic role in managing standard-density workloads, particularly when paired with direct liquid cooling or rear-door heat exchangers. In this context, their cost-effective deployment, especially in regions with favourable climates, ensures that air-based systems will remain part of the cooling portfolio for years to come, albeit in a narrower and more targeted operational scope.

Liquid-based cooling systems: commercial maturity and future trajectory

The rising demand for AI, HPC, and GPU-intensive applications has intensified thermal challenges across the data centre sector, driving a transition from conventional air-based cooling to more efficient liquid-based systems. Offering significantly improved heat transfer, lower energy consumption, and enhanced potential for heat reuse, liquid cooling is no longer confined to research or niche deployments. While broad market penetration is still evolving, commercial adoption is accelerating, particularly in hyperscale cloud platforms and specialised AI clusters.

Among the available technologies, DLC has emerged as the most commercially mature approach. It relies on cold plates attached directly to high-power components (i.e., CPUs and GPUs), circulating coolant through closed loops to remove heat at the source. Major operators, including Equinix, Digital Realty, Microsoft, and CoolIT have deployed DLC across select data centres, motivated by its ability to support higher rack densities and reduce operational energy use [246]. Reports indicate DLC can lower cooling-related energy consumption by up to 72.4 %, translating into significant energy savings and emissions reductions

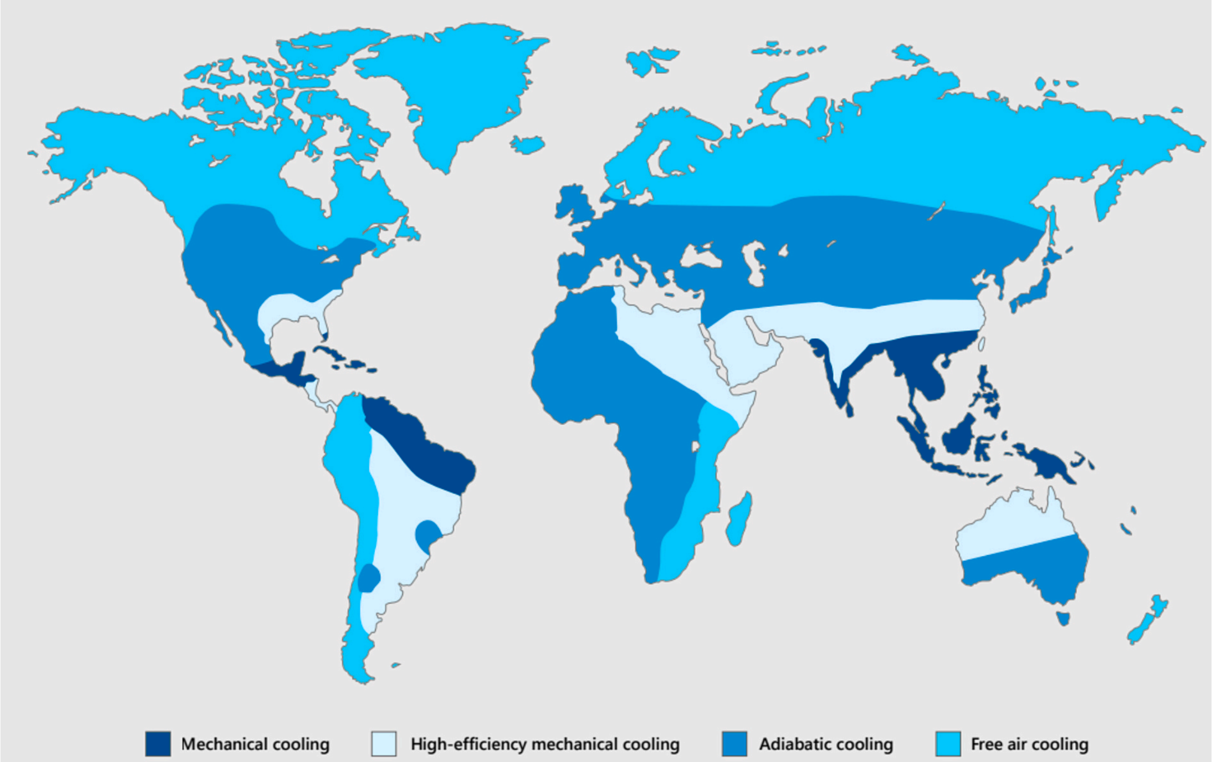

Fig. 5. Data centre cooling methods by climate region. Adopted from [242].

[247].

The economic case for DLC is strengthening. Analysis by CoolIT suggests potential annual savings of approximately $259,000/MW of IT heat load. At the GWDG research facility in Germany, DLC deployment raised rack power densities from 23.5 kW to 96 kW and improved PUE from 1.24 to 1.07 [247]. DLC also enables thermal integration with district heating, as demonstrated at Stockholm's atNorth facility, reinforcing its appeal under ESG and sustainability frameworks.

In parallel, technological innovations are improving DLC scalability. Microsoft's custom "sidekick" system combines DLC with microfluidic cooling, embedding narrow liquid channels directly into chips, allowing precise thermal management with minimal water use [248]. Operators such as Equinix have adopted hybrid layouts where DLC systems coexist with air-based systems, enabling gradual and modular transitions within existing infrastructure [249].

Despite these advantages, barriers to widespread DLC adoption remain. Maintenance complexity and high initial implementation costs are cited by over 50 % of IT professionals, while nearly half identify a lack of internal expertise as a significant constraint [246]. Additional concerns include risks of leaks/spills, the recurring cost of coolant replacement, and system design requirements such as warm-water loop support and fluid containment infrastructure. These challenges explain why more than 60 % of organisations are projected to remain reliant on air cooling by 2026, with only 6.5 % adopting immersion-based hybrid configurations, and less than 2 % expected to transition to full immersion cooling [250].

While DLC is scaling, immersion cooling, in which servers are submerged in dielectric fluids, offers even greater thermal performance, supporting rack densities above 100 kW and PUE as low as 1.03 [251]. Although still limited to purpose-built environments, immersion cooling has seen experimental deployments by operators such as Equinix and through partnerships involving Submer, Intel, and Shell. This approach delivers up to 95 % energy savings and 90 % water savings but demands significant architectural redesign, including custom enclosures, reinforced floors, and advanced fluid management systems [252]. The chemistry and lifecycle cost of immersion fluids also play a critical role in commercial viability, particularly in mission-critical environments or remote edge deployments.

Strategic investments from companies like Schneider Electric, NVIDIA, Lenovo, and HPE reflect growing industry confidence in both DLC and immersion cooling. As summarised in Table 4, recent product launches and acquisitions are enabling broader adoption across AI and HPC use cases. These developments confirm that liquid cooling is no longer experimental, it is becoming essential for thermally intensive workloads and next-generation infrastructure.

Cooling technologies beyond DLC and immersion

While DLC and immersion cooling represent the current commercial frontier, spray cooling and jet impingement cooling are emerging as promising technologies for specialised deployment contexts. These innovations aim to extend the flexibility and scalability of liquid-based systems, particularly in compact, modular, or high-ambient environments.

Spray cooling uses fine, atomised droplets of dielectric fluid to cool components directly, enabling heat removal without bulky heatsinks or complex loops. Though still in pilot stages, companies like PMC Technologies have developed oil-based spray systems for sealed edge modules, while NTU Singapore demonstrated PUE levels of 1.08 using dielectric spray systems in tropical climates [264,265]. Spray cooling offers low-noise, fanless operation and is well-suited for edge computing, ruggedised modules, and water-constrained regions.

Jet impingement cooling, by contrast, uses high-velocity fluid jets to achieve extremely high heat transfer rates at the chip level. While not yet commercially widespread, collaborations (i.e., Danfoss and HyprCool) have produced sealed, modular chassis with integrated impingement spray systems, tailored for AI inference nodes and mobile data centres [266,267]. These systems eliminate airflow reliance and may eventually complement DLC and immersion in niche, high-intensity environments.

Hybrid cooling systems market adoption

Hybrid cooling systems have become a key component in the evolving thermal management strategies of modern data centres. As operators face growing pressure to accommodate high-density compute

Table 4
Summary of key developments in liquid cooling technology for AI and HPC data centres.

Company	Key Development	Technology Focus	Strategic Impact	Source
Schneider Electric and Motivair Corp.	Acquired 75 % stake in Motivair for up to $850 M	DLC, High-Capacity Thermal Systems	Expands Schneider's portfolio to address thermal demands of AI and HPC workloads	[253]
NVIDIA	Integrated advanced liquid cooling into GB200 NVL72 AI server systems	Liquid-cooled AI infrastructure with Blind Mate Liquid Cooling Manifold and Floating Tray Connection	Supports ultra-high-density AI compute, helping reduce energy use in facilities projected to consume 8 % of U.S. grid electricity by 2030	[254]
Supermicro	Launched DLC-enabled clusters for hyperscale data centres	High-performance DLC	Boosts performance, reduces cooling energy costs, and supports green data centres	[255]
Hewlett Packard Enterprise (HPE)	Introduced AMD-based systems with optional DLC for AI and HPC	Hybrid air–liquid cooling	Offers scalable and sustainable compute solutions for enterprise AI workloads	[256]
iDataCool (IBM and University of Regensburg)	Advanced research on hot water cooling and waste heat reuse	Hot-water liquid cooling, thermal energy recovery	Enhances energy reuse and sustainability in academic and industrial HPC centres	[257]
Boyd	Developed plug-and-play liquid cooling for NVIDIA GB200 NVL72	Full liquid cooling system integration	Accelerates deployment and improves energy efficiency of NVIDIA-based AI servers	[258]
Lenovo	Released ThinkSystem N1380 Neptune Chassis and SC777 V4 Neptune server	Liquid-cooled AI servers with NVIDIA Blackwell GPUs	Enables 100 % heat removal for racks exceeding 100 kW without the need for traditional air cooling	[259]
Intel and Submer	Co-developed immersion-cooled Intel Xeon processors, FCHS package for single-phase cooling	Single-phase immersion cooling	Enables > 1000 W thermal design power (TDP) chip cooling; improves sustainability, supports dense AI/HPC systems	[260]
Submer	Developed SmartPod platform for immersion cooling	Immersion cooling (140 kW per rack), biodegradable fluid	Achieves PUE between 1.03 and 1.1; supports high-density workloads	[261]
Shell	Introduced gas-to-liquid (GTL)-based immersion cooling fluids; deployed GRC tanks in Houston data centre	Immersion cooling fluids	Supports eco-friendly, biodegradable, and thermally stable cooling systems	[262]
Sandia National Labs and Submer	Testing Submer's immersion system in HPC cluster	Immersion cooling with non-conductive liquid	Potential to cut power use by 70 % in HPC-AI clusters, and could reduce 100 % of the generated heat	[263]

while managing costs and legacy constraints, hybrid architectures provide a pragmatic middle ground, enabling phased transitions to liquid cooling without abandoning existing air-based infrastructure. Their growing adoption reflects both technological flexibility and a strong commercial case for staged investment. The appeal of hybrid systems lies in their ability to incrementally scale thermal performance, especially in facilities where a full transition to liquid cooling is economically or operationally infeasible. Rather than relying solely on chilled air or advanced liquid systems, hybrid deployments combine elements of both to balance capital expenditure, rack density requirements, and facility lifecycle planning.

In commercial practice, the most prominent form of hybrid cooling today is the RDHx. While the technical characteristics of RDHx systems have been discussed in Section 4.1.2, their commercial relevance is increasingly evident. Leading colocation and hyperscale providers, Equinix, Digital Realty, ColdLogik, and HPE, have adopted RDHx as a retrofit-friendly pathway to support AI and HPC workloads. These systems are now marketed not only for their thermal performance, but also for their role in preserving existing white space, avoiding hot aisle reconfiguration, and minimising infrastructure disruption during deployment [268].

Financially, RDHx systems offer a strong value proposition. Studies suggest energy savings of up to 90 % compared to traditional CRAC-based systems, particularly by reducing the need for room-wide air circulation and enabling localised, rack-level heat extraction [269]. In retrofit scenarios, they eliminate the cost and downtime associated with full liquid retrofits or immersion tank installations. Additionally, integration with DLC systems, such as those marketed by Motivair and HPE, further enhances their commercial appeal, enabling scalable hybrid configurations that evolve with workload demands [270].

Beyond RDHx, newer hybrid platforms such as Vertiv’s CoolPhase Flex introduce dynamic cooling modes that shift between air and liquid operation depending on real-time thermal loads. These systems incorporate AI-driven control algorithms, economiser-based free cooling, and adaptive liquid engagement to optimise energy use across seasonal and workload variability [271]. Such capabilities are particularly appealing in regions with fluctuating ambient conditions or unpredictable compute cycles.

Equinix’s Cool Array represents another example of hybrid innovation tailored for high-humidity environments. Designed for use in tropical climates like Singapore, the system integrates fan-wall arrays and cold-flooded containment to reduce PUE while maintaining airflow integrity in dense rack layouts [249]. Equinix has demonstrated that hybrid designs can maintain thermal compliance under diverse climatic and regulatory contexts by selectively augmenting air-cooling infrastructure with modular enhancements.

From a strategic perspective, hybrid cooling enables phased investment strategies that align with IT refresh cycles, ESG reporting goals, and operational risk management. Operators frequently begin with RDHx to enable rack-level cooling gains, integrate DLC in high-density zones, and reserve immersion systems for mission-specific clusters [272]. This modular upgrade model avoids stranded assets and allows for cooling architectures to evolve in tandem with compute demand, particularly important as AI and edge workloads become increasingly variable in thermal profile and geographic distribution.

However, barriers to broader hybrid adoption remain. Facilities must accommodate chilled water routing, CDU integration, and airflow management modifications. Moreover, maintaining component interoperability across hybrid systems requires adherence to open standards such as Open19 and ASHRAE guidelines, especially as multi-vendor deployments become more common [249]. These challenges underscore the need for robust planning and systems integration expertise during hybrid implementation.

Table 5 compares key cooling technologies, including air, liquid, and hybrid options, across deployment conditions, adopters, and business considerations to situate hybrid systems within the broader commercial landscape. Hybrid systems stand out for their deployment flexibility, cost efficiency, and retrofit compatibility, making them a natural default in facilities preparing for higher rack densities without committing to full immersion or DLC.

Future trends and emerging technologies

AI-driven cooling

As global data processing demand increases, efficient cooling becomes critical for data centres. AI-driven cooling technologies offer transformative potential by optimising thermal management, energy efficiency, and predictive maintenance. AI systems analyse historical data and real-time sensor inputs to detect early signs of equipment degradation, enabling proactive repairs, reducing downtime, and lowering operational costs. Enhanced fault detection also speeds up troubleshooting, minimising operational disruption.

These AI-driven systems operate across three functional layers: sensor data collection, AI-based control and optimisation, and actuation of cooling equipment. Fig. 6 illustrates this integrated architecture, showing how real-time data from sensors informs predictive control logic, which dynamically adjusts cooling operations and enables heat reuse pathways. As depicted in the central layer, AI modules perform forecasting, optimisation, and control functions to manage complex thermal dynamics. In the upper control tier of the figure, actuators adjust fan and pump speeds, regulate liquid flow, and redirect waste heat to reuse applications.

When integrated with traditional cooling techniques, AI enables real-time adaptability and efficiency gains. For example, hybrid surrogate models merging Artificial Neural Networks (ANNs) with thermofluid equations can predict temperature and pressure distributions with a Root Mean Square Error (RMSE) of 0.52 [281]. These models outperform CFD simulations in speed, delivering results in under 4 s, and are now used for thermally aware workload management and cooling configurations optimisation, reducing the need for physical testing and computational overhead while boosting energy efficiency.

Beyond hybrid surrogate models, AI-enhanced evaporative cooling has shown notable energy efficiency gains. In one study, water mist evaporation in an overhead downward flow system reduced air temperature from 21.63 °C to 18 °C, increased humidity from 50 % to 70 %, and achieved a PUE of 1.42, below the global average [90]. AI techniques like Gaussian Process Regression and Reinforcement Learning enable real-time control, dynamically adjusting system parameters to optimise performance while preventing condensation, making these systems both efficient and adaptive.

Beyond evaporative systems, rack-level cooling technologies such as loop thermosyphons (LTS) showcase the synergy between advanced cooling and AI for improved efficiency and fault resilience. A study demonstrated that LTS recovered 93.8 % of cooling capacity during a fan failure by optimising backup airflow, cold water temperatures, and flow rate [282]. Incorporating AI to manage these parameters in real time could further improve system resilience and energy efficiency, solidifying LTS as a promising solution for next generation data centres.

Complementing rack-level approaches, row-based cooling units deliver precise temperature control with minimal energy consumption by optimising airflow configurations. A zonal transient model based on mechanical resistances achieves accurate temperature predictions with reduced computational complexity compared to CFD or standalone machine learning models. This approach cut cooling power consumption by 22 %, using two cooling units instead of one [283]. Integrating AI frameworks such as Reinforcement Learning or Gaussian Process Regression could further enhance real-time control of airflow, water temperature, and server loads.

Researchers are increasingly combining CFD simulations with AI models to improve both accuracy and adaptability in data centre cooling. A multi-scale hybrid framework integrating CFD with a CNN-

Table 5
Comparative overview of commercial data centre cooling technologies.

Cooling Type	Technology	Used By (Examples)	Deployment Conditions	Thermal Efficiency	Scalability (kW/rack)	Environmental Impact	Business Impact	Barriers and Risks	Source
Air-based	CRAC	Equinix, Telehouse, Hyve	Small/medium data centres (<200 kW); low-to-moderate rack densities (<15 kW)	PUE typically ranges from 1.5 to 2.0, reflecting relatively low cooling efficiency	Supports up to ~ 25–40 kW/rack with airflow containment	No direct water usage; limited heat reuse potential; average exhaust air ~ 30 °C	Low-cost, simple-to-integrate cooling system for legacy facilities	Poor performance at higher rack densities; rising energy costs from inefficient fan use	[240,245,273]
	CRAH	Equinix, Telehouse	Large-scale facilities (>200 kW); requires chilled water plant	Higher efficiency than CRAC; PUE ranges from 1.2 to 1.5	Supports up to ~ 25–40 kW/rack with optimised airflow	Water consumption can reach ~ 25.5 M litres/MW/year; WUE ~ 1.8 L/kWh; limited heat reuse (~30 °C)	Suitable for facilities with chilled water infrastructure; scalable for large loads	Complex setup; not viable in all climates; expensive for small deployments	[240,241,274]
	Free Air Cooling (Direct)	Microsoft, Google, AWS	Cold climates (e.g., Scandinavia, Canada); low humidity regions using filtered outdoor air for direct cooling	PUE as low as 1.15; can reduce cooling energy consumption by 40–90 %	Typically supports 10–25 kW/rack	Near-zero water use; heat reuse feasible (exhaust 30–50 °C)	Excellent sustainability and cost profile in favourable climates	Not viable in humid or polluted environments; depends on seasonal variability	[242,275,276]
	Free Air Cooling (Indirect)			PUE ~ 1.2–1.4; ~18 % energy savings compared to adiabatic systems	Typically supports 25–40 kW/rack	Water use ~ 8.4 M L/MW/year; better air quality control; heat reuse between 25–45 °C	Preferred in regions with air quality concerns; isolates IT space	Slight efficiency penalty from heat transfer losses	[277,278]
	Evaporative / Adiabatic	Microsoft, AWS, Google	Common in dry or temperate zones; supplements air-side cooling with misting or wetted pads	Improves partial pPUE by ~ 0.01 compared to standard air-cooled CRAC systems	Up to ~ 40 kW/rack in well-optimised deployments	Water use ranges from 0.5 M to 8.4 M litres/MW/year depending on climate and method; lower heat reuse potential	Reduces energy cost in arid environments; balances cooling with moderate water consumption	Water sourcing and treatment are critical; unsuitable for humid climates	[241,242,277]
Liquid-based	Direct Liquid Cooling (DLC)	Equinix, Digital Realty, Microsoft, CoolIT, Lenovo	Applied in AI/HPC clusters and dense compute environments where rack power exceeds air cooling limits	PUE between 1.07 and 1.1; heat capture typically 70–75 %; energy reduction up to 72.4 %	Supports densities up to 300 + kW/rack with cold plates or microchannel systems	Closed-loop systems eliminate direct water loss; warm water output (~40–45 °C) enables integration with district heating and industrial reuse	Reduces OPEX by ~$259 K/MW annually; supports ESG targets via energy savings and heat reuse	High CapEx; requires CDUs, fluid monitoring, and facility retrofitting	[247,250]
	Immersion Cooling	Submer, Intel, Shell, Sandia Labs, Equinix (pilot)	Deployed in purpose-built HPC and AI centres, components submerged in dielectric fluid	PUE as low as 1.03; up to 40 % cooling energy reduction; near-complete heat capture (~95–100 %)	Supports 200–368 kW/rack depending on cooling fluid and setup	Eliminates cooling towers; ~55 °C heat output suitable for high-grade reuse (district heating, greenhouses, industry)	Highest thermal efficiency and water savings; suitable for extreme-density deployments	High infrastructure costs; redesign of racks and tanks; fluid handling and compatibility concerns	[261,263,273]
Hybrid	Rear-Door Heat Exchangers (RDHx)	Equinix, Digital Realty, ColdLogik, HPE, Motivair	Deployed in retrofit-friendly scenarios where rear-door cooling integrates with air-cooled setups	Cooling energy PUE between 1.25 and 1.33; reduces cooling energy by 30–40 %; heat output ~ 40 °C	Typically supports 40–100 kW/rack; some configurations up to 200 kW/rack	WUE ranges from 0.01 to 0.635 L/kWh, depending on temperature difference; supports low-grade heat reuse	Allows gradual transition to liquid cooling; minimises disruption in existing facilities	Requires CDU setup, pipe routing, trained personnel; transient response tuning needed	[270,279,280]
	CoolPhase Flex (Dynamic Hybrid)	Vertiv, Compass Datacenters	Flexible workloads; transitions from air to	Variable; supports free cooling in low-load	Adaptive; transitions from standard	Designed for water minimisation; uses dry cooling	AI-optimised; ideal for phasing between legacy	New architecture; higher initial cost; reliant on	[271]

(continued on next page)

Table 5 (*continued*)

Cooling Type	Technology	Used By (Examples)	Deployment Conditions	Thermal Efficiency	Scalability (kW/rack)	Environmental Impact	Business Impact	Barriers and Risks	Source
			liquid operation based on thermal load	states; efficiency driven by predictive control models	to high-density (air + liquid)	or misting as needed	and next-gen systems	software control performance	
	Cool Array	Equinix	Tropical and high-humidity climates (e.g., Singapore); designed for optimised airflow	Enhances airflow with high-efficiency fan walls; improves PUE over legacy air cooling	Effective up to moderate densities; not ideal beyond ~ 30–40 kW/rack	Reduces fan energy use; no direct water use; limited heat reuse	Deployed across over 30 Equinix sites; improves energy performance in warm regions	Less suitable for ultra-high-density workloads; limited beyond air-cooled boundaries	[249]

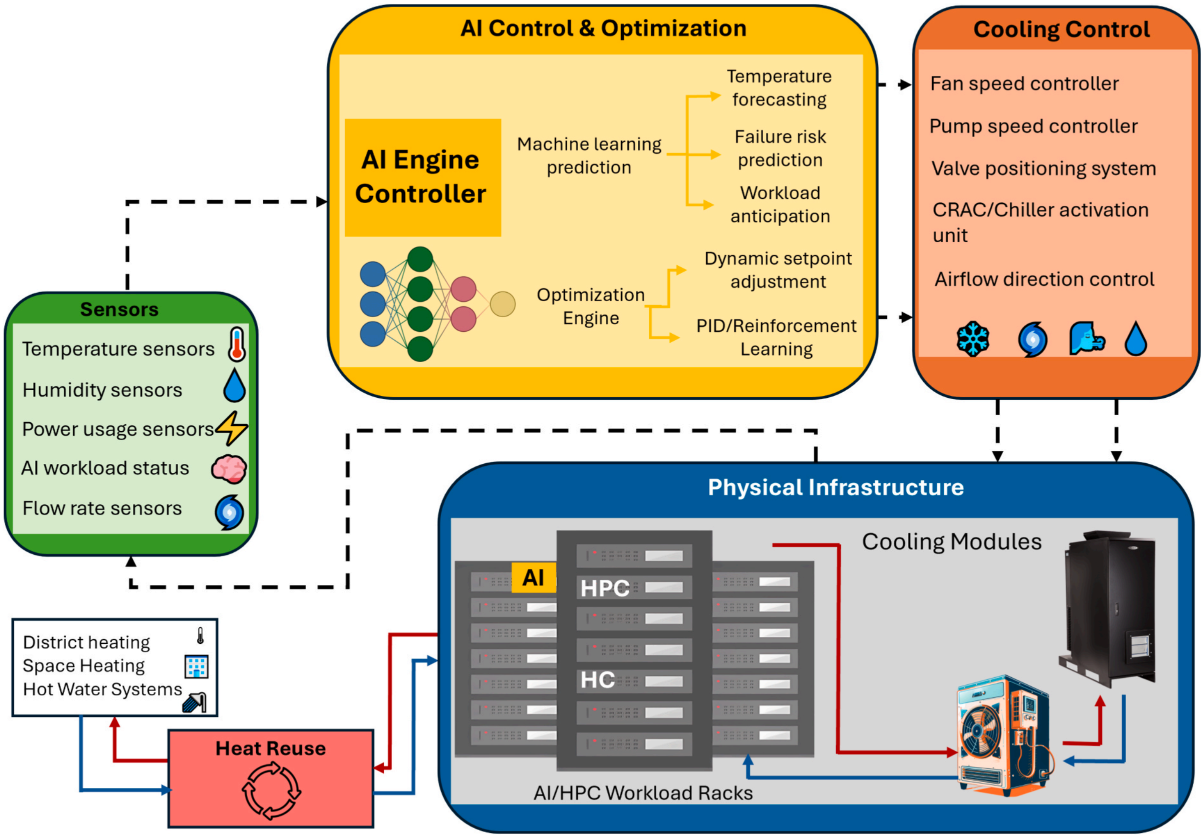

Fig. 6. AI-driven data centre cooling framework.

BiLSTM-Attention model achieved a prediction accuracy of $R^2 = 0.9899$, reducing CPU temperatures by 12.47 K as airflow increased from 4 m/s to 6 m/s [284]. Incorporating Bayesian optimisation further refined predictions and enabled real-time adaptive cooling based on workload and environmental changes [285].

Within this ecosystem, Rack Mountable Cooling Units (RMCUs) and Federated Learning offer scalable and privacy-preserving approaches to optimise energy efficiency. RMCUs allow for precise, rack-level airflow control in decentralised systems. A parameter-free transient zonal model achieved real-time temperature predictions with accuracy within 4 % of experimental results, while outperforming CFD in computational efficiency [286]. Integrating AI into these systems could support dynamic airflow adjustments and IT load distribution, further enhancing energy efficiency and system adaptability across distributed architectures.

Recent advancements (i.e., Multimodal Federated Learning (MMFL)-based thermal-aware job scheduling) highlight AI's potential for improving energy efficiency. Utilising multimodal data inputs such as server heatmaps and workload data, models like CNN and XGBoost reduced thermal imbalances by 20 % [287]. Additionally, federated learning enables privacy-preserving optimisation across distributed facilities.

AI-driven cooling strategies also benefit from model predictive control (MPC) [288]. An MPC framework for rack-based cooling achieved a 9.7 % energy savings compared to traditional PID controllers by adjusting cooling parameters based on real-time server workloads [289]. Advanced AI techniques like Reinforcement Learning could further enhance MPC adaptability in nonlinear conditions.

In chilled water systems, AI-powered adaptive control algorithms are driving efficiency. For instance, ANN-based controllers accurately predicted CRAH supply air temperatures and optimised chilled water flow rates [290]. Among tested retraining techniques, the sliding window method delivered the best performance with an RMSE of 0.08 °C. These predictive controls help maintain thermal stability while minimising energy consumption in fluctuating IT environments.

Finally, grey-box models that combine thermodynamic principles with data-driven approaches provide accurate and computationally efficient temperature predictions for real-time applications. One model enabled significant energy savings by dynamically adjusting airflow rates and IT loads based on workload variations [291]. A state-space model designed for rack-based architectures further demonstrated the value of combining deterministic and AI-driven methods to support sustainable data centre operations [292].

Emerging technologies

The data centre industry is undergoing a significant transformation, driven by advancements in computational efficiency and thermal management. As infrastructure complexity increases, emerging technologies are key to enhancing sustainability, performance, and reliability. Fig. 7 illustrates how these advancements are shaping the future of data centres.

While Fig. 7 offers a qualitative visualisation, the progression of each technology is informed by several underlying criteria: (i) its technology readiness level (TRL) based on recent demonstration studies; (ii) the extent of peer-reviewed and industry literature indicating feasibility and validation; (iii) current commercial availability or pilot deployment status; and (iv) integration complexity with existing data centre infrastructure. Technologies such as hybrid AI-cooling systems and heat reuse architectures are positioned further along the timeline due to demonstrated case studies and market adoption trends. In contrast, quantum-augmented optimisation and bio-inspired cooling are earlier-stage innovations with significant potential but limited real-world validation. This roadmap serves as a directional overview to situate these technologies within the broader data centre evolution context.

Federated learning (FL) enables decentralised AI training across distributed facilities, optimising cooling systems while preserving data privacy. The application of FL in geo-distributed environments is still limited, but as workloads increase, traditional job scheduling algorithms become less effective. FL allows model parameters, not raw data, to be shared, enabling predictive thermal control and hotspots across preventing while complying with data privacy regulations [287]. Integrating multimodal approaches, (i.e., combining heatmap images with workload data) further enhances predictive accuracy and operational efficiency [293].

While FL is widely adopted in sectors like healthcare and finance, its application in power systems remains still emerging. Studies have used FL to analyse smart meter data, predict electricity demand, and classify customer loads with privacy protection [294]. However, its role in energy aggregation, particularly in forecasting energy demand between Energy Aggregation Service Providers (EASPs) and suppliers, remains unexplored. Additionally, in data centres, FL supports energy-efficient cooling, allowing real-time, distributed optimisation of thermal performance while reducing computational overhead [295].

AI-driven predictive cooling networks build upon utilising digital twins and reinforcement learning to dynamically adjust thermal performance in real-time. Advanced models of convolutional neural networks (CNNs), long short-term memory (LSTM) networks, and graph neural networks (GNNs) optimise cooling responses to workload and environmental variations [296]. Digital twins facilitate real-time monitoring, while Gaussian Process Regression and deep reinforcement learning refine temperature predictions, improving energy consumption and thermal stability. Studies indicate these systems can reduce power consumption by up to 22 % compared to traditional cooling methods [297].

Meanwhile, innovative materials are reshaping thermal management. Graphene-based and liquid metal materials provide superior thermal conductivity and energy storage, reducing reliance on active cooling solutions. Self-healing systems with smart pipelines detect and repair leaks autonomously. PCMs help manage thermal spikes, reducing hotspots by 6–10 % and absorbing 94.55 % of incoming power during the melting phase [298]. However, their low post-melt thermal conductivity limits long-term heat dissipation. Ongoing research into PCM/metal heat exchangers and hybrid integration with loop heat pipes or cold plates has demonstrated 46–66 % power savings by decreasing reliance on traditional fans and water pumps [299].

Thermal energy reuse is another promising sustainability, repurposing data centre waste heat for district heating, agricultural, and energy generation. This aligns with global energy efficiency goals by reintegrating excess heat into local ecosystems. Meanwhile, quantum computing, while offering transformative potential for solving complex optimisation problems, introduces significant cooling challenges. Unlike conventional systems, quantum data centres often consume more energy for cryogenic cooling than for the computation itself [300]. Cooling demands vary depending on qubit architecture, packaging, and the integration of cryogenic and room-temperature circuits. First-principles energy models suggest that reducing cryogenic power is essential for sustainable quantum infrastructure.

On the optimisation side, quantum algorithms are being explored to manage energy in AI data centres. The Variational Quantum Computing-based Robust Optimisation (VQC-RO) framework integrates quantum circuits with classical optimisation for uncertainty-aware control, achieving 9.8 % lower power consumption and 12.5 % lower carbon emissions [301]. Similarly, hybrid quantum–classical Benders' decomposition algorithms have demonstrated up to 40 % fewer iterations than classical methods in large-scale data centre energy management, illustrating the growing role of quantum tools in sustainable operations [302].

Looking ahead, bio-inspired cooling technologies are emerging as promising solutions for energy-efficient data centres. Biomimetic heat exchangers, modelled on vascular networks, optimise fluid distribution

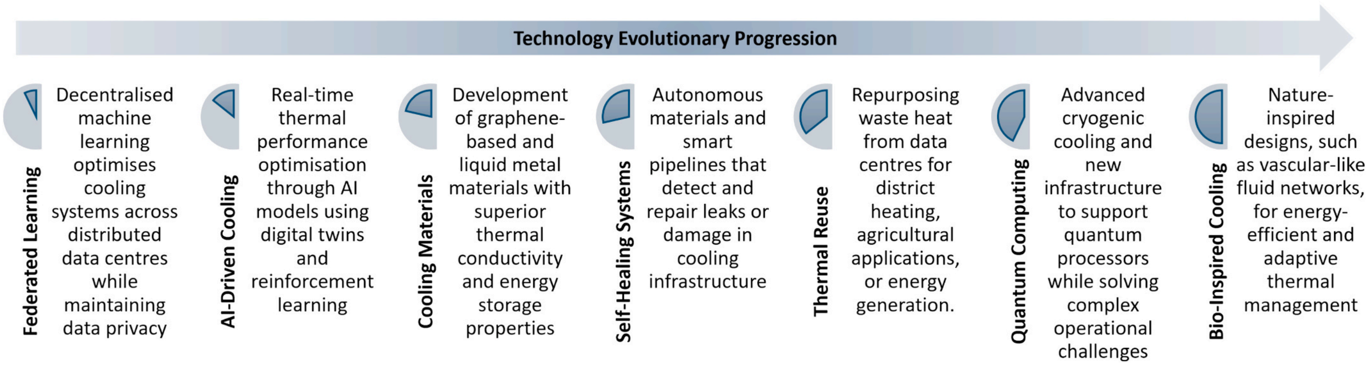

Fig. 7. Visually represents the interplay between these technologies and their impact on future data centre developments.

for effective thermal management [303]. Designs inspired by termite mounds enable passive airflow regulation, reducing reliance on mechanical cooling, while superhydrophobic coatings, mimicking lotus leaves, improve liquid cooling efficiency by minimising adhesion and thermal resistance [304]. These nature-inspired techniques offer sustainable solutions for reducing energy consumption while improving thermal management efficiency.

Thermoelectric cooling also represents a promising avenue in emerging thermal management research. Based on the Peltier effect, these solid-state systems can provide localised, precision cooling at the component level without requiring moving parts or fluid-based loops [305]. This makes them particularly attractive for modular edge data centres, compact AI inference nodes, and sealed environments where mechanical complexity and maintenance must be minimised. Although current thermoelectric materials face efficiency limitations for high-power workloads, ongoing research into AI-optimised thermoelectric control and next-generation semiconductor materials may enable their future integration as part of low-noise, scalable, and environmentally friendly cooling architectures [306].

Collectively, these emerging technologies, ranging from AI-driven optimisations and thermoelectric cooling to advanced materials and bio-inspired designs, represent the next frontier in data centre evolution. Their success will require ongoing research, industry collaboration, and strategic investments to overcome deployment barriers and scale adoption.

Conclusion

The convergence of AI, HPC, and hyperscale computing is redefining the thermal profile of modern data centres, rendering traditional air-based cooling architectures increasingly inadequate. This review has evaluated the evolution of cooling technologies, with a particular focus on their commercial readiness, deployment conditions, and strategic relevance in the context of rising compute intensity, regional constraints, and sustainability mandates.

Among the available solutions, DLC and RDHx have emerged as the most commercially mature and operationally viable options. These systems offer a clear return on investment, compatibility with existing infrastructure, and phased deployment models that allow organisations to scale with workload demands. DLC systems, especially cold plate configurations, are now standard in AI-driven hyperscale facilities, supporting densities beyond 50 kW per rack while enabling heat reuse. RDHx systems, by contrast, are particularly effective in enterprise and colocation environments, where modular retrofits and capital efficiency are essential.

Although immersion cooling offers industry-leading thermal efficiency and PUE values as low as 1.03, it remains commercially niche, limited to purpose-built AI and HPC deployments. Barriers such as architectural redesign, fluid handling complexity, and safety concerns have slowed its broader uptake. However, recent innovations in biodegradable and GTL-based fluids, sealed plug-and-play systems, and vendor-neutral chassis standards suggest that immersion cooling could become more viable in the longer term, particularly for high-TDP edge workloads or in sustainability-committed greenfield developments.

Hybrid cooling systems are playing a pivotal role in enabling this transition. By combining elements of air and liquid cooling, they allow operators to extend the life of existing infrastructure while preparing for next-generation densities. RDHx and hybrid DLC systems (i.e., those deployed by Vertiv, ColdLogik, and Equinix) offer a practical and economically scalable roadmap for modernising thermal management without the need for wholesale infrastructure changes.

Geographic and environmental factors further influence technology selection. In water-stressed regions, operators should prioritise free air cooling, indirect evaporative systems, and warm-water DLC, avoiding open-loop adiabatic cooling unless reclaimed or closed-loop systems are in place. In colder climates, operators should capitalise on free cooling and district heat reuse, aligning cooling investments with broader ESG and circular economy goals.

Standardisation will be essential to ensuring future compatibility, resilience, and auditability of cooling systems. Frameworks from ASHRAE, the Open Compute Project (OCP), and the Energy Efficient High-Performance Computing (EEHPC) working group are already shaping reference architectures and safety protocols. Widespread adoption of open formats such as Open19 V2, along with support for negative pressure cooling loops and automatic de-energisation, will play a central role in reducing risk, streamlining procurement, and supporting multi-vendor interoperability.

Looking forward, flexibility and modularity will define successful thermal strategies. A strategic approach begins with RDHx, progresses through DLC integration, and evolves toward immersion or sealed systems as workload intensity increases. This phased investment model reduces upfront cost, limits operational disruption, and allows infrastructure to adapt alongside changing compute requirements. In parallel, AI-driven thermal optimisation, real-time predictive control, and energy-aware scheduling will further reduce energy consumption and improve system resilience. The integration of waste heat reuse, particularly through district heating networks or on-site industrial use, offers not only environmental benefits but also new value streams.

Ultimately, liquid and hybrid cooling are no longer speculative or emergent, they are now central to the business and operational models of future-ready data centres. Organisations that adopt standards-compliant, modular, and ESG-aligned cooling strategies will be best positioned to lead in both technological innovation and sustainable performance. Despite growing adoption, several areas require further investigation to accelerate the next wave of cooling innovation. These include:

1. Long-term reliability and safety testing of immersion fluids and negative-pressure systems in production environments.
2. Economic modelling of phased retrofits in mixed-density facilities.
3. Real-time AI-based control integration for thermal optimisation and dynamic workload allocation.
4. In-depth environmental impact studies, including ESG performance metrics and lifecycle CO_2 assessment of liquid vs. air-based systems.
5. Scalability of waste heat recovery, particularly in multi-tenant or colocation contexts.
6. Regulatory frameworks for global standardisation of next-generation cooling infrastructure.

These research priorities will be instrumental in guiding data centre operators, technology providers, and policymakers as they navigate the intersection of performance, sustainability, and digital growth in the AI era.

CRediT authorship contribution statement

Dlzar Al Kez: Writing – original draft, Visualization, Investigation, Conceptualization. **Aoife M Foley:** Writing – review & editing, Project administration, Funding acquisition, Conceptualization. **Fadhli Wong B.M. Hasan Wong:** Writing – review & editing, Project administration, Funding acquisition. **Andrea Dolfi:** Writing – review & editing, Project administration, Funding acquisition. **Geetha Srinivasan:** Writing – review & editing, Project administration, Funding acquisition.

Declaration of competing interest

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

Acknowledgements

The authors gratefully acknowledge the funding support from PETRONAS. This research is also funded by the UKRI's Engineering and Physical Sciences Research Council (EPSRC) under grant EP/X038823/2. The views and opinions expressed in this work may not necessarily reflect those of PETRONAS, UKRI's Engineering, and EPSRC.

Data availability

No data was used for the research described in the article.

References

[1] Fortune Business Insights, "Smart Strategies, Giving Speed to your Growth Trajectory," 2025. https://www.fortunebusinessinsights.com/data-center-market-109851 (accessed Jan. 30, 2025).

[2] Synergy Research Group, "Hyperscale Data Centers Hit the Thousand Mark; Total Capacity is Doubling Every Four Years," 2024. https://www.srgresearch.com/articles/hyperscale-data-centers-hit-the-thousand-mark-total-capacity-is-doubling-every-four-years (accessed Jan. 30, 2025).

[3] J. Aljbour, T. Wilson, and Patel Poorvi, Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption, California, USA, 2024. [Online]. Available: https://www.epri.com/research/products/000000003002028905.

[4] International Energy Agency (IEA), "Electricity 2024," Paris, France, 2024. [Online]. Available: https://www.iea.org/reports/electricity-2024.

[5] Al Kez D, Foley AM. Exploring the sustainability challenges facing digitalization and internet data centers. J Clean Prod 2022;371:133633. https://doi.org/10.1016/j.jclepro.2022.133633.

[6] de Vries A. The growing energy footprint of artificial intelligence. Joule 2023;7(10):2191–4. https://doi.org/10.1016/j.joule.2023.09.004.

[7] S. Kimball, Microsoft signs deal to invest more than $10 billion on renewable energy capacity to power data centers, *CNBC*, 2024. https://www.cnbc.com/2024/05/01/microsoft-brookfield-to-develop-more-than-10point5-gigawatts-of-renewable-energy.html (accessed Jan. 30, 2025).

[8] Derrick M. Digital Realty & schneider: powering net zero data centres. Digital Real 2024. https://energydigital.com/articles/digital-realty-expands-renewable-energy-offering-to-1-4gw (accessed Jan. 30, 2025).

[9] Yuan X, Liang Y, Hu X, Xu Y, Chen Y, Kosonen R. Waste heat recoveries in data centers: a review. Renew Sustain Energy Rev 2023;188:113777. https://doi.org/10.1016/j.rser.2023.113777.

[10] Isazadeh A, Ziviani D, Claridge DE. Thermal management in legacy air-cooled data centers: an overview and perspectives. Renew Sustain Energy Rev 2023;187:113707. https://doi.org/10.1016/j.rser.2023.113707.

[11] Cai S, Gou Z. Towards energy-efficient data centers: a comprehensive review of passive and active cooling strategies. Energy Built Environ 2024. https://doi.org/10.1016/j.enbenv.2024.08.009.

[12] Ni J, Bai X. A review of air conditioning energy performance in data centers. Renew Sustain Energy Rev 2017;67:625–40. https://doi.org/10.1016/j.rser.2016.09.050.

[13] Silva CA, Vilaça R, Pereira A, Bessa RJ. A review on the decarbonization of high-performance computing centers. Renew Sustain Energy Rev 2024;189:114019. https://doi.org/10.1016/j.rser.2023.114019.

[14] Riu I, Smiley D, Bessasparis S, Patel K. Load growth is here to stay, but are data centers? Strategically Manag Challeng Opportunities Load Growth 2024.

[15] Pramanik S. Electronic Cooler Technologies and Superior Data Center Cooling Techniques. IGI Global Scientific Publishing; 2025. p. 26. https://doi.org/10.4018/979-8-3693-5563-3.ch012.

[16] Ramakrishnan B, et al. Understanding the impact of data center liquid cooling on energy and performance of machine learning and artificial intelligence workloads. J Electron Packag 2025;147(2). https://doi.org/10.1115/1.4067136.

[17] CTM Magnetics, "Air Cooled vs. Liquid Cooled: The Differences," 2025. https://ctmmagnetics.com/general/air-cooled-vs-liquid-cooled-differences/#:~:text=It is striking to see,3%2C500 times that of air. (accessed Jan. 30, 2025).

[18] Heydari A, et al. Experimental evaluation of direct-to-chip cold plate liquid cooling for high-heat-density data centers. Appl Therm Eng 2024;239:122122. https://doi.org/10.1016/j.applthermaleng.2023.122122.

[19] Heydari A, et al. Parameters of performance: a deep dive into liquid-to-air CDU assessment. Case Stud Therm Eng 2024;64:105461. https://doi.org/10.1016/j.csite.2024.105461.

[20] Birbarah P, et al. Water immersion cooling of high power density electronics. Int J Heat Mass Transf 2020;147:118918. https://doi.org/10.1016/j.ijheatmasstransfer.2019.118918.

[21] Roe C, et al. Immersion cooling for lithium-ion batteries – a review. J Power Sources 2022;525:231094. https://doi.org/10.1016/j.jpowsour.2022.231094.

[22] Dede EM, et al. Techno-economic feasibility analysis of an extreme heat flux micro-cooler. iScience 2023;26(1):105812. https://doi.org/10.1016/j.isci.2022.105812.

[23] Shah JM, Padmanaban K, Singh H, Asokan SD, Saini S, Agonafer D. Evaluating the reliability of passive server components for single-phase immersion cooling. J Electron Packag 2022;144(2). https://doi.org/10.1115/1.4052536.

[24] Williams NP, Trimble D, Shaughnessy SMO. Thermal performance of lithium-ion battery tabs under liquid immersion cooling conditions. Futur Batter 2025;5:100037. https://doi.org/10.1016/j.fub.2025.100037.

[25] Pambudi NA, Sarifudin A, Firdaus RA, Ulfa DK, Gandidi IM, Romadhon R. The immersion cooling technology: current and future development in energy saving. Alexandria Eng J 2022;61(12):9509–27. https://doi.org/10.1016/j.aej.2022.02.059.

[26] Zhou F, Gu W, Ma G. Advancements in data center cooling systems: from refrigeration to high performance cooling. Energ Buildings 2024;320:114634. https://doi.org/10.1016/j.enbuild.2024.114634.

[27] ISO/IEC, "Information technology, Data centres, Key performance indicators, Part 2: Power usage effectiveness (PUE)," 2016. [Online]. Available: https://cdn.standards.iteh.ai/samples/63451/87c396de9e76423088236bf87cdf73c1/ISO-IEC-30134-2-2016.pdf.

[28] Li Y, Zhu C, Li X, Yang B. A review of non-uniform load distribution and solutions in data centers: micro-scale liquid cooling and large-scale air cooling. Energies 2025;18(1). https://doi.org/10.3390/en18010149.

[29] Azarifar M, Arik M, Chang JY. Liquid cooling of data centers: a necessity facing challenges. Appl Therm Eng 2024;247:123112. https://doi.org/10.1016/j.applthermaleng.2024.123112.

[30] Li Z, et al. Comprehensive review and future prospects on chip-scale thermal management: core of data center's thermal management. Appl Therm Eng 2024;251:123612. https://doi.org/10.1016/j.applthermaleng.2024.123612.

[31] Ngarambe J, Yun GY, Santamouris M. The use of artificial intelligence (AI) methods in the prediction of thermal comfort in buildings: energy implications of AI-based thermal comfort controls. Energ Build 2020;211:109807. https://doi.org/10.1016/j.enbuild.2020.109807.

[32] Foley AM, Heffron RJ, Al Kez D, Furszyfer Del Rio DD, McInerney C, Welfle A. Restoring trust in ESG investing through the adoption of just transition ethics. Renew Sustain Energy Rev 2024;199:114557. https://doi.org/10.1016/j.rser.2024.114557.

[33] I. Seaton, "Cooling Efficiency Algorithms: Economizers and Temperature Differentials (Water-Side Economizers – Parallel)," *Upsite Technologies*, 2019. https://www.upsite.com/blog/cooling-efficiency-algorithms-economizers-and-temperature-differentials-water-side-economizers-parallel/ (accessed Apr. 22, 2025).

[34] Tang Y, Bai X, Jin C, Fu Y, Mao W, Xu X. Experimental and numerical study on the thermal environment of a data center. Build Environ 2022;212:108741. https://doi.org/10.1016/j.buildenv.2021.108741.

[35] Zhang Y, et al. Experimental study on the performance of a novel data center air conditioner combining air cooling and evaporative cooling. Int J Refrig 2025;170:98–112. https://doi.org/10.1016/j.ijrefrig.2024.11.010.

[36] Zhang K, Zhang Y, Liu J, Niu X. Recent advancements on thermal management and evaluation for data centers. Appl Therm Eng 2018;142:215–31. https://doi.org/10.1016/j.applthermaleng.2018.07.004.

[37] Cho J, Kim Y. Development of modular air containment system: thermal performance optimization of row-based cooling for high-density data centers. Energy 2021;231:120838. https://doi.org/10.1016/j.energy.2021.120838.

[38] Lu H, Zhang Z. Numerical and experimental investigations on the thermal performance of a data center. Appl Therm Eng 2020;180:115759. https://doi.org/10.1016/j.applthermaleng.2020.115759.

[39] Isazadeh A, Ziviani D, Claridge DE. Global trends, performance metrics, and energy reduction measures in datacom facilities. Renew Sustain Energy Rev 2023;174:113149. https://doi.org/10.1016/j.rser.2023.113149.

[40] Verified Market Reports, "CRAC and CRAH Units Market Insights," 2025. https://www.verifiedmarketreports.com/product/crac-and-crah-units-market/ (accessed Mar. 13, 2025).

[41] Li L, Jin C, Bai X. A determination method on the supply airflow rate with the impact of leakage and bypass airflow of computer room air conditioning. J Build Eng 2022;59:105063. https://doi.org/10.1016/j.jobe.2022.105063.

[42] Sun K, Luo N, Luo X, Hong T. Prototype energy models for data centers. Energ Build 2021;231:110603. https://doi.org/10.1016/j.enbuild.2020.110603.

[43] Zhao R, Du Y, Yang X, Zhou Z, Wang W, Yang X. A critical review on the thermal management of data center for local hotspot elimination. Energ Build 2023;297:113486. https://doi.org/10.1016/j.enbuild.2023.113486.

[44] Manaserh YM, Tradat MI, Gharaibeh AR, Sammakia BG, Tipton R. Shifting to energy efficient hybrid cooled data centers using novel embedded floor tiles heat exchangers. Energy Convers Manag 2021;247:114762. https://doi.org/10.1016/j.enconman.2021.114762.

[45] Zou Z, Yu X, Ergan S. Towards optimal control of air handling units using deep reinforcement learning and recurrent neural network. Build Environ 2020;168:106535. https://doi.org/10.1016/j.buildenv.2019.106535.

[46] Fulpagare Y, Hsu PH, Wang CC. Experimental analysis of airflow uniformity and energy consumption in data centers. Appl Therm Eng 2022;209:118302. https://doi.org/10.1016/j.applthermaleng.2022.118302.

[47] Salih Erden H. Comprehensive energy and economic assessment of CRAH bypass method in air-cooled data centers. Sustain Energy Technol Assess 2022;52:102120. https://doi.org/10.1016/j.seta.2022.102120.

[48] Cho J, Woo J. Development and experimental study of an independent row-based cooling system for improving thermal performance of a data center. Appl Therm Eng 2020;169:114857. https://doi.org/10.1016/j.applthermaleng.2019.114857.

[49] Zou S, Zhang Q, Yue C. Comparative study on different energy-saving plans using water-side economizer to retrofit the computer room air conditioning system. J Build Eng 2023;69:106278. https://doi.org/10.1016/j.jobe.2023.106278.

[50] Ahmadi VE, Erden HS. A parametric CFD study of computer room air handling bypass in air-cooled data centers. Appl Therm Eng 2020;166:114685. https://doi.org/10.1016/j.applthermaleng.2019.114685.

[51] Chen H, et al. Numerical investigation of thermal performance with adaptive terminal devices for cold aisle containment in data centers. Buildings 2023;13(2). https://doi.org/10.3390/buildings13020268.

[52] Cool Shield, Aisle Containment Systems: Build-to-Suit Aisle Containment Solutions by CoolShield. https://cool-shield.com/aisle-containment/ (accessed Mar. 25, 2025).

[53] Gupta R, Asgari S, Moazamigoodarzi H, Pal S, Puri IK. Cooling architecture selection for air-cooled data centers by minimizing exergy destruction. Energy 2020;201:117625. https://doi.org/10.1016/j.energy.2020.117625.

[54] Wibron E, Ljung AL, Staffan Lundström T. Comparing performance metrics of partial aisle containments in hard floor and raised floor data centers using CFD. Energies 2019;12(8). https://doi.org/10.3390/en12081473.

[55] Wang F, Huang Y, Prasetyo BY. Energy-efficient improvement approaches through numerical simulation and field measurement for a data center. Energies 2019;12(14). https://doi.org/10.3390/en12142757.

[56] Cho J, Woo J, Park B, Lim T. A comparative CFD study of two air distribution systems with hot aisle containment in high-density data centers. Energies 2020;13(22). https://doi.org/10.3390/en13226147.

[57] Jiang D. Effects and optimization of airflow on the thermal environment in a data center. Front Built Environ 2024;10:1–12. https://doi.org/10.3389/fbuil.2024.1362861.

[58] Chu WX, Wang R, Hsu PH, Wang CC. Assessment on rack intake flowrate uniformity of data center with cold aisle containment configuration. J Build Eng 2020;30:101331. https://doi.org/10.1016/j.jobe.2020.101331.

[59] Ni J, Jin B, Ning S, Wang X. The numerical simulation of the airflow distribution and energy efficiency in data centers with three types of aisle layout. Sustain 2019;11(18). https://doi.org/10.3390/su11184937.

[60] Aili A, Long W, Cao Z, Wen Y. Radiative free cooling for energy and water saving in data centers. Appl Energy 2024;359:122672. https://doi.org/10.1016/j.apenergy.2024.122672.

[61] Silva-Llanca L, Ponce C, Bermúdez E, Martínez D, Díaz AJ, Aguirre F. Improving energy and water consumption of a data center via air free-cooling economization: the effect weather on its performance. Energy Convers Manag 2023;292:117344. https://doi.org/10.1016/j.enconman.2023.117344.

[62] Borkowski M, Piłat AK. Customized data center cooling system operating at significant outdoor temperature fluctuations. Appl Energy 2022;306:117975. https://doi.org/10.1016/j.apenergy.2021.117975.

[63] Zhou Y, et al. Energy savings in direct air-side free cooling data centers: a cross-system modeling and optimization framework. Energ Build 2024;308:114003. https://doi.org/10.1016/j.enbuild.2024.114003.

[64] Lan J, Zhang Z, Liang X, Wu H, Wang G, Mao R. Experimental and numerical investigation on thermal performance of data center via fan-wall free cooling technology. Appl Therm Eng 2023;228:120467. https://doi.org/10.1016/j.applthermaleng.2023.120467.

[65] Zhang J, Mao R, Li C, Lan J, Yi X, Zhang Z. Optimization air-conditioning system and thermal management of data center via fan-wall free cooling technology. Appl Therm Eng 2023;234:121245. https://doi.org/10.1016/j.applthermaleng.2023.121245.

[66] Van Le D, Liu Y, Wang R, Tan R, Ngoh LH. Air free-cooled tropical data center: design, evaluation, and learned lessons. IEEE Trans Sustain Comput 2022;7(3):579–94. https://doi.org/10.1109/TSUSC.2021.3132927.

[67] Swegon, High-efficiency solutions for data centres, 2024. https://www.swegon.com/uk/knowledge-hub/application-guides/data-centres/ (accessed Feb. 01, 2025).

[68] Chen X, Wang X, Wang L, Zheng H, Ding T, Li Z. Multistage data center cooling system for temperature gradation and matching. Appl Energy 2025;377:124592. https://doi.org/10.1016/j.apenergy.2024.124592.

[69] Van Zetten JD, Cholette ME, Bamdad K. Data centre building energy optimization with differential temperature control utilizing I.T. power measurement. J Build Eng 2025;100:111722. https://doi.org/10.1016/j.jobe.2024.111722.

[70] Kim JH, Shin DU, Kim H. Data center energy evaluation tool development and analysis of power usage effectiveness with different economizer types in various climate zones. Buildings 2024;14(1). https://doi.org/10.3390/buildings14010299.

[71] Wang Z, Zeng C, Zhu Z, Li Y, Ma X, Zhao X. Time-series machine learning for predictive optimisation of a highly efficient evaporative cooling system. Build Serv Eng Res Technol 2025:1–14. https://doi.org/10.1177/01436244251315047.

[72] Hnayno M, Chehade A, Klaba H, Polidori G, Maalouf C. Experimental investigation of an optimized indirect free cooling system including a dry cooler equipped with evaporative cooling pads for data center. Energy Rep 2023;9:460–9. https://doi.org/10.1016/j.egyr.2023.12.005.

[73] 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 2023;171:112991. https://doi.org/10.1016/j.rser.2022.112991.

[74] Guo P, Wang S, Lei Y, Li J. Numerical simulation of solar chimney-based direct airside free cooling system for green data centers. J Build Eng 2020;32:101793. https://doi.org/10.1016/j.jobe.2020.101793.

[75] Ljungqvist HM, Risberg M, Toffolo A, Vesterlund M. A realistic view on heat reuse from direct free air-cooled data centres. Energy Convers Manag X 2023;20:100473. https://doi.org/10.1016/j.ecmx.2023.100473.

[76] Chu J, Huang X. Research status and development trends of evaporative cooling air-conditioning technology in data centers. Energy Built Environ 2023;4(1):86–110. https://doi.org/10.1016/j.enbenv.2021.08.004.

[77] Han Z, Sun X, Wei H, Ji Q, Xue D. Energy saving analysis of evaporative cooling composite air conditioning system for data centers. Appl Therm Eng 2021;186:116506. https://doi.org/10.1016/j.applthermaleng.2020.116506.

[78] Chang Q, Huang Y, Liu K, Xu X, Zhao Y, Pan S. Optimization control strategies and evaluation metrics of cooling systems in data centers. A Review", Sustain. 2024;16(16). https://doi.org/10.3390/su16167222.

[79] Lee YT, Wen CY, Shih YC, Li Z, Yang AS. Numerical and experimental investigations on thermal management for data center with cold aisle containment configuration. Appl Energy 2022;307:118213. https://doi.org/10.1016/j.apenergy.2021.118213.

[80] De Antonellis S, Joppolo CM, Liberati P, Milani S, Romano F. Modeling and experimental study of an indirect evaporative cooler. Energ Build 2017;142:147–57. https://doi.org/10.1016/j.enbuild.2017.02.057.

[81] Zhang Y, et al. Cooling technologies for data centres and telecommunication base stations – a comprehensive review. J Clean Prod 2022;334:130280. https://doi.org/10.1016/j.jclepro.2021.130280.

[82] Cui X, Yang C, Yan W, Wang X, Liu Y, Jin L. Climatic applicability of indirect evaporative cooling strategies for data centers in China. J Build Eng 2024;83:108431. https://doi.org/10.1016/j.jobe.2023.108431.

[83] Lionello M, Lucchese R, Rampazzo M, Guay M, Atta K. Energy-efficient operation of indirect adiabatic data center cooling systems via Newton-like phasor extremum seeking control. Int J Adapt Control Signal Process 2021;35(7):1298–315. https://doi.org/10.1002/acs.3223.

[84] Jo MS, Shin JH, Kim WJ, Jeong JW. Energy-saving benefits of adiabatic humidification in the air conditioning systems of semiconductor cleanrooms. Energies 2017;10(11):1–23. https://doi.org/10.3390/en10111774.

[85] Yan W, et al. Multi-objective optimization of dew point indirect evaporative coolers for data centers. Appl Therm Eng 2024;241:122425. https://doi.org/10.1016/j.applthermaleng.2024.122425.

[86] Oxycom, Two-stage evaporative cooling. https://www.oxy-com.com/two-stage-evaporative-cooling (accessed Mar. 26, 2025).

[87] Hwang J. Daytime radiative cooling under extreme weather conditions. Adv. Energy Sustain. Res. 2024;5(5). https://doi.org/10.1002/aesr.202300239.

[88] Borodinecs A, et al. Feasibility of reducing electricity consumption of air conditioning equipment by condenser direct evaporative cooling technology. Example of case study in Dubai. Atmosphere (Basel) 2021;12(9). https://doi.org/10.3390/atmos12091205.

[89] Han Z, et al. Simulation study on performance of data center air-conditioning system with novel evaporative condenser. Energy 2020;210:118521. https://doi.org/10.1016/j.energy.2020.118521.

[90] Lee YC, Chen KY, Yan WM, Shih YC, Chao CY. Evaporative cooling method to improve energy management of overhead downward flow-type data center. Case Stud Therm Eng 2023;45:102998. https://doi.org/10.1016/j.csite.2023.102998.

[91] Marques GC, von Seggern F, Dehm S, Breitung B, Hahn H, Dasgupta S. Influence of humidity on the performance of composite polymer electrolyte-gated field-effect transistors and circuits. IEEE Trans Electron Devices 2019;66(5):2202–7. https://doi.org/10.1109/TED.2019.2903456.

[92] Bergfried C, Abdi Qezeljeh S, Roisman IV, De Gersem H, Hussong J, Späck-Leigsnering Y. Thermal finite-element model of electric machine cooled by spray. Energies 2025;18(1):1–16. https://doi.org/10.3390/en18010084.

[93] Li C, Mao R, Wang Y, Zhang J, Lan J, Zhang Z. Experimental study on direct evaporative cooling for free cooling of data centers. Energy 2024;288:129889. https://doi.org/10.1016/j.energy.2023.129889.

[94] Ndukaife TA, Nnanna AGA. Optimization of water consumption in hybrid evaporative cooling air conditioning systems for data center cooling applications. Heat Transf Eng 2019;40(7):559–73. https://doi.org/10.1080/01457632.2018.1436418.

[95] Ndukaife TA, Nnanna AGA. Enhancement of performance and energy efficiency of air conditioning system using evaporatively cooled condensers. Heat Transf Eng 2019;40(3–4):375–87. https://doi.org/10.1080/01457632.2018.1429063.

[96] Yan W, et al. Breaking the limits of ambient air conditions in evaporative cooling with vacuum-assisted hollow fiber membrane technology. Energ Buildings 2024;319:114557. https://doi.org/10.1016/j.enbuild.2024.114557.

[97] Aili A, Yin X, Yang R. Passive sub-ambient cooling: radiative cooling versus evaporative cooling. Appl Therm Eng 2022;202:117909. https://doi.org/10.1016/j.applthermaleng.2021.117909.

[98] AIRSYS Cooling Technologies, Evaporative Cooling for Data Centers — Pros and Cons, 2024. https://airsysnorthamerica.com/evaporative-cooling-for-data-centers-pros-and-cons/ (accessed Jan. 26, 2025).

[99] Habibi Khalaj A, Halgamuge SK. A Review on efficient thermal management of air- and liquid-cooled data centers: from chip to the cooling system. Appl Energy 2017;205:1165–88. https://doi.org/10.1016/j.apenergy.2017.08.037.

[100] Wonder Group, Five Ways To Cooling Blade Servers, 2022. https://cpuheatsink.com/blog/five-ways-to-cooling-blade-servers/ (accessed Jan. 21, 2025).

[101] Batson A, Thorpe D, Duncan G. Global Data Center Outlook: Shaping tomorrow's digital infrastructure. London: United Kingdom; 2025.

[102] Pambudi S, Saechua W, Jongyingcharoen JS. A thermogravimetric assessment of eco-friendly biochar from oxidative torrefaction of spent coffee grounds: combustion behavior, kinetic parameters, and potential emissions. Environ Technol Innov 2024;33:103472. https://doi.org/10.1016/j.eti.2023.103472.

[103] Chen H, Li D. Current status and challenges for liquid-cooled data centers. Front Energy Res 2022;10:1–4. https://doi.org/10.3389/fenrg.2022.952680.

[104] Zhang Y, Li C, Pan M. Design and performance research of integrated indirect liquid cooling system for rack server. Int J Therm Sci 2023;184:107951. https://doi.org/10.1016/j.ijthermalsci.2022.107951.

[105] Zhang Q, et al. A survey on data center cooling systems: technology, power consumption modeling and control strategy optimization. J Syst Archit 2021;119:102253. https://doi.org/10.1016/j.sysarc.2021.102253.

[106] Zhou X, Xin Z, Tang W, Sheng K, Wu Z. Energy, economic and environmental analysis of a combined cooling, power generation, and energy storage system: a case study of data center in Shenzhen. J Energy Storage 2024;97:112968. https://doi.org/10.1016/j.est.2024.112968.

[107] Zhang H, Ma H, Ma S, Yang M. Investigation on the performance of an indirect evaporative cooling system integrated with liquid dehumidification. Energ Build 2021;251:111356. https://doi.org/10.1016/j.enbuild.2021.111356.

[108] He W, et al. Performance optimization of server water cooling system based on minimum energy consumption analysis. Appl Energy 2021;303:117620. https://doi.org/10.1016/j.apenergy.2021.117620.

[109] He W, et al. Optimal thermal management of server cooling system based cooling tower under different ambient temperatures. Appl Therm Eng 2022;207:118176. https://doi.org/10.1016/j.applthermaleng.2022.118176.

[110] Lamaison N, Ong CL, Marcinichen JB, Thome JR. Two-phase mini-thermosyphon electronics cooling: dynamic modeling, experimental validation and application to 2U servers. Appl Therm Eng 2017;110:481–94. https://doi.org/10.1016/j.applthermaleng.2016.08.198.

[111] Yu ZQ, Li MT, Cao BY. A comprehensive review on microchannel heat sinks for electronics cooling. Int J Extrem Manuf 2024;6(2). https://doi.org/10.1088/2631-7990/ad12d4.

[112] Zhou J, Lu M, Zhao Q, Li Q, Chen X. Topologically optimized manifold microchannel heat sink with extreme cooling performance for high heat flux cooling of electronics. Appl Therm Eng 2024;241:122426. https://doi.org/10.1016/j.applthermaleng.2024.122426.

[113] Wu T, Wang L, Tang Y, Yin C, Li X. Flow and heat transfer performances of liquid metal based microchannel heat sinks under high temperature conditions. Micromachines 2022;13(1). https://doi.org/10.3390/mi13010095.

[114] Das L, Habib K, Saidur R, Aslfattahi N, Yahya SM, Rubbi F. Improved thermophysical properties and energy efficiency of aqueous ionic liquid/mxene nanofluid in a hybrid pv/t solar system. Nanomaterials 2020;10(7):1–26. https://doi.org/10.3390/nano10071372.

[115] Japar WMAA, Sidik NAC, Saidur R, Asako Y, Nurul Akmal Yusof S. A review of passive methods in microchannel heat sink application through advanced geometric structure and nanofluids: current advancements and challenges. Nanotechnol Rev 2020;9(1):1192–216.

[116] Zhang J, Wu J, Xie Z, Lu Z, Guan X, Ge Y. Suitability of embedded liquid cooling and heat generation for chips. Micromachines 2024;15(1). https://doi.org/10.3390/mi15010009.

[117] Sadiq Al-Baghdadi MAR, Noor ZMH, Zeiny A, Burns A, Wen D. CFD analysis of a nanofluid-based microchannel heat sink. Therm Sci Eng Prog 2020;20:100685. https://doi.org/10.1016/j.tsep.2020.100685.

[118] Alkrush AA, Salem MS, Abdelrehim O, Hegazi AA. Data centers cooling: a critical review of techniques, challenges, and energy saving solutions. Int J Refrig 2023;160:246–62. https://doi.org/10.1016/j.ijrefrig.2024.02.007.

[119] Hassani SM, Khoshvaght-Aliabadi M, Mazloumi SH, Rehman S, Alimoradi A. Improving thermal performance of microchannels by combining rectangular pin with chamber. Appl Therm Eng 2021;186:116373. https://doi.org/10.1016/j.applthermaleng.2020.116373.

[120] Alugoju UK, Dubey SK, Javed A. Optimization of converging and diverging microchannel heat sink for electronic chip cooling. IEEE Trans Components, Packag Manuf Technol 2020;10(5):817–27. https://doi.org/10.1109/TCPMT.2020.2985402.

[121] Zhang M, et al. Liquid metal manifold microchannel heat sink for ultra-high heat flux cooling. Appl Therm Eng 2024;248(PA):123117. https://doi.org/10.1016/j.applthermaleng.2024.123117.

[122] Yang S, Li J, Cao B, Wu Z, Sheng K. Investigation of Z-type manifold microchannel cooling for ultra-high heat flux dissipation in power electronic devices. Int J Heat Mass Transf 2024;218:124792. https://doi.org/10.1016/j.ijheatmasstransfer.2023.124792.

[123] Lucchese R, Varagnolo D, Johansson A. Controlled direct liquid cooling of data servers. IEEE Trans Control Syst Technol 2021;29(6):2325–38. https://doi.org/10.1109/TCST.2019.2942270.

[124] Gullbrand J, Luckeroth MJ, Sprenger ME, Winkel C. Liquid cooling of compute system. J. Electron. Packag. 2019;141(1). https://doi.org/10.1115/1.4042802.

[125] W. Homberg, C. Eric, B. Genci, R. J. Psnc, N. M. Psnc, and D. P. Jsc, Ways to Decrease TCO and Infrastructure Related Costs, 2024. Accessed: Jan. 21, 2025. [Online]. Available: https://prace-ri.eu/wp-content/uploads/WP262.pdf.

[126] Alkasmoul FS, Asaker M, Almogbel A, AlSuwailem A. Combined effect of thermal and hydraulic performance of different nanofluids on their cooling efficiency in microchannel heat sink. Case Stud Therm Eng 2022;30:101776. https://doi.org/10.1016/j.csite.2022.101776.

[127] Karimi L, et al. Water-energy tradeoffs in data centers: a case study in hot-arid climates. Resour Conserv Recycl 2022;181:106194. https://doi.org/10.1016/j.resconrec.2022.106194.

[128] Siddik MAB, Shehabi A, Marston L. The environmental footprint of data centers in the United States. Environ. Res. Lett. 2021;16(6). https://doi.org/10.1088/1748-9326/abfba1.

[129] Zhao C, Jing Y, Xie X, Jiang Y. Application and simulation of IEC water chillers for summer free cooling in data centers. Appl Therm Eng 2025;260:124915. https://doi.org/10.1016/j.applthermaleng.2024.124915.

[130] Jing Y, et al. Field test of cooling systems in two air-cooled data centers: various regions, air distributions and evaporative cooling technologies. Appl Therm Eng 2024;248:123189. https://doi.org/10.1016/j.applthermaleng.2024.123189.

[131] Shi W, Ma X, Min Y, Yang H. Feasibility analysis of indirect evaporative cooling system assisted by liquid desiccant for data centers in hot-humid regions. Sustain. 2024;16(5). https://doi.org/10.3390/su16052011.

[132] Bux C, Rana RL, Lombardi M, Giungato P, Tricase C. A critical analysis of global warming potential of data centers in the digital era. Int J Life Cycle Assess 2025. https://doi.org/10.1007/s11367-024-02419-2.

[133] Wang C, Yu H, Xian L, Liu G, Chen L, Tao WQ. Study on the operation strategy of the water cooling system in the single-phase immersion cooling data center considering cooling tower performance. Energ Buildings 2024;324. https://doi.org/10.1016/j.enbuild.2024.114890.

[134] Ding B, Xia F, He L, Huang Y. A novel concept for performance enhancement of immersion-cooled data center servers. Int Commun Heat Mass Transf 2024;159:107980. https://doi.org/10.1016/j.icheatmasstransfer.2024.107980.

[135] Sun X, Liu Z, Ji S, Yuan K. Experimental study on thermal performance of a single-phase immersion cooling unit for high-density computing power data center. Int J Heat Fluid Flow 2025;112:109735. https://doi.org/10.1016/j.ijheatfluidflow.2024.109735.

[136] Liu C, Huang Y, Zhang C. Efficient jet-assisted single-phase immersion liquid cooling for high heat-flux servers. Appl Therm Eng 2025;259. https://doi.org/10.1016/j.applthermaleng.2024.124935.

[137] Lin Y, Luo Y, Li W, Cao Y, Tao Z, Shih TIP. Single-phase and two-phase flow and heat transfer in microchannel heat sink with various manifold arrangements. Int J Heat Mass Transf 2021;171:121118. https://doi.org/10.1016/j.ijheatmasstransfer.2021.121118.

[138] Cheng CC, Chang PC, Li HC, Hsu FI. Design of a single-phase immersion cooling system through experimental and numerical analysis. Int J Heat Mass Transf 2020;160:120203. https://doi.org/10.1016/j.ijheatmasstransfer.2020.120203.

[139] Abo-Zahhad EM, et al. Flow boiling in a four-compartment heat sink for high-heat flux cooling: a parametric study. Energy Convers Manag 2021;230:113778. https://doi.org/10.1016/j.enconman.2020.113778.

[140] 3M Company, "3M to Exit PFAS Manufacturing by the End of 2025," 2022. https://news.3m.com/2022-12-20-3M-to-Exit-PFAS-Manufacturing-by-the-End-of-2025 (accessed Mar. 28, 2025).

[141] Engineering Update, "A range of drop-in replacements for 3M^{TM} Novec^{TM} Engineered Fluids from EnviroTech Europe," 2024. https://engineering-update.co.uk/2024/09/27/a-range-of-drop-in-replacements-for-3m-novec-engineered-fluids-from-envirotech-europe/ (accessed Mar. 28, 2025).

[142] T. Wodoslawsky, Is 'Fatal Heat' Murdering Your Components? Consider Liquid Cooling, Beave, USA, 2024. [Online]. Available: https://img.c3controls.com/c3controls-Consider-Liquid-Cooling-2023-0525.pdf.

[143] The Chemours Company, Opteon^{TM} 2P50: Developmental Dielectric Heat Transfer Fluid, Singapore, 2024. [Online]. Available: https://www.opteon.com/en/-/media/files/opteon/opteon-2p50-gic-product-info-sheet-fnl-100523.pdf?rev=1b708201ecd7489b8a9cf86c4de7ffc5.

[144] Liu C, et al. Experimental investigation on oil spray cooling with hairpin windings. IEEE Trans Ind Electron 2020;67(9):7343–53. https://doi.org/10.1109/TIE.2019.2942563.

[145] Henshaw K, Cooper P. Single-phase vs. two-phase Immersion cooling: what's the difference and which one is right for you? Submer 2021. https://submer.com/blog/single-phase-vs-two-phase-immersion-cooling/ (accessed Feb. 03, 2025).

[146] Fung B, Beran L. Achieving Sustainable Data Center Growth. USA: Redwood City; 2022.

[147] Wahab A, Najmi AUH, Senobar H, Amjady N, Kemper H, Khayyam H. Immersion cooling innovations and critical hurdles in Li-ion battery cooling for future electric vehicles. Renew Sustain Energy Rev 2024;211:115268. https://doi.org/10.1016/j.rser.2024.115268.

[148] Liu C, Yu H. Evaluation and optimization of a two-phase liquid-immersion cooling system for data centers. Energies 2021;14(5). https://doi.org/10.3390/en14051395.

[149] Fazli M, Mehrjardi SAA, Mahmoudi A, Khademi A, Amini M. Advancements in pulsating heat pipes: exploring channel geometry and characteristics for enhanced thermal performance. Int J Thermofluids 2024;22:100644. https://doi.org/10.1016/j.ijft.2024.100644.

[150] Zhang C, Sun X, Han Z, Li X, Dong J. Energy saving potential analysis of two-phase immersion cooling system with multi-mode condenser. Appl Therm Eng 2023;219:119614. https://doi.org/10.1016/j.applthermaleng.2022.119614.

[151] LiquidStack, Two-Phase vs Single-Phase Immersion Cooling Fluids: Deconstructing Myths with Science, Carrollton, USA, 2022.

[152] F2 Chemicals Ltd, Electrical Properties of Perfluorocarbons, 2022. https://www.f2chemicals.com/pdf/technical/Electrical prop.pdf (accessed Jan. 02, 2025).

[153] Owens J, Xiao A, Bonk J, DeLorme M, Zhang A. Recent development of two alternative gases to SF6 for high voltage electrical power applications. Energies 2021;14(16):1–4. https://doi.org/10.3390/en14165051.

[154] Salierno G. On the chemical pathways influencing the effective global warming potential of commercial hydrofluoroolefin gases. ChemSusChem 2024;17(17). https://doi.org/10.1002/cssc.202400280.

[155] Open Compute Project, Material Compatibility in Immersion Cooling, Austin, USA, 2022.

[156] Tong Z, Han Z, Fang C, Wen X. Two-phase thermosyphon loop with different working fluids used in data centers. Int J Heat Mass Transf 2023;214:124393. https://doi.org/10.1016/j.ijheatmasstransfer.2023.124393.

[157] Wu X, et al. Investigations on heat dissipation performance and overall characteristics of two-phase liquid immersion cooling systems for data center. Int

J Heat Mass Transf 2025;239:126575. https://doi.org/10.1016/j.ijheatmasstransfer.2024.126575.

[158] Liu C, Yu H. Experimental investigations on heat transfer characteristics of direct contact liquid cooling for CPU. Buildings 2022;12(7). https://doi.org/10.3390/buildings12070913.

[159] Sajjad U, Sadeghianjahromi A, Ali HM, Wang CC. Enhanced pool boiling of dielectric and highly wetting liquids - a review on enhancement mechanisms. Int Commun Heat Mass Transf 2020;119:104950. https://doi.org/10.1016/j.icheatmasstransfer.2020.104950.

[160] Yuan X, et al. Phase change cooling in data centers: a review. Energ Buildings 2021;236. https://doi.org/10.1016/j.enbuild.2021.110764.

[161] Udaya Kumar G, Suresh S, Thansekhar MR, Halpati D. Role of inter-nanowire distance in metal nanowires on pool boiling heat transfer characteristics. J Colloid Interface Sci 2018;532:218–30. https://doi.org/10.1016/j.jcis.2018.07.092.

[162] Gong Y, Zhou F, Ma G, Liu S. Advancements on mechanically driven two-phase cooling loop systems for data center free cooling. Int J Refrig 2022;138:84–96. https://doi.org/10.1016/j.ijrefrig.2022.03.007.

[163] Kanbur BB, Wu C, Fan S, Duan F. System-level experimental investigations of the direct immersion cooling data center units with thermodynamic and thermoeconomic assessments. Energy 2021;217:119373. https://doi.org/10.1016/j.energy.2020.119373.

[164] Sun X, Han Z, Li X. Simulation study on cooling effect of two-phase liquid-immersion cabinet in data center. Appl Therm Eng 2021;207:118142. https://doi.org/10.1016/j.applthermaleng.2022.118142.

[165] Zhou G, Zhou J, Huai X, Zhou F, Jiang Y. A two-phase liquid immersion cooling strategy utilizing vapor chamber heat spreader for data center servers. Appl Therm Eng 2022;210:118289. https://doi.org/10.1016/j.applthermaleng.2022.118289.

[166] Haghshenas K, Setz B, Blosch Y, Aiello M. Enough hot air: the role of immersion cooling. Energy Informatics 2023;6(1). https://doi.org/10.1186/s42162-023-00269-0.

[167] Ramakrishnan B, et al. CPU overclocking: a performance assessment of air, cold plates, and two-phase immersion cooling. IEEE Trans Components, Packag Manuf Technol 2021;11(10):1703–15. https://doi.org/10.1109/TCPMT.2021.3106026.

[168] Mody F, Chauhan A, Shukla M, Kandlikar SG. Evaluation of heater size and external enhancement techniques in pool boiling heat transfer with dielectric fluids. Int J Heat Mass Transf 2022;183:122176. https://doi.org/10.1016/j.ijheatmasstransfer.2021.122176.

[169] Amorim GS, Belman-Flores JM, de Paoli Mendes R, Sandoval OR, Khosravi A, Garcia-Pabon JJ. Recent advancements in thermal management technologies for cooling of data centers. J Brazilian Soc Mech Sci Eng 2024;46(8):1–20. https://doi.org/10.1007/s40430-024-05048-w.

[170] Wang H, Gan Y, Li R, Liu F, Li Y. Experimental study on the thermal performance of a liquid-cooled heat sink integrating heat pipes for dual CPU servers. Appl Therm Eng 2024;236:121851. https://doi.org/10.1016/j.applthermaleng.2023.121851.

[171] She X, et al. Thermodynamic analysis for new solutions of data center cooling with adsorption refrigeration and heat pipe. Int J Refrig 2024;160:341–56. https://doi.org/10.1016/j.ijrefrig.2024.01.016.

[172] Sun X, et al. Experimental study on a novel pump-driven heat pipe/vapor compression system for rack-level cooling of data centers. Energy 2023;274:127335. https://doi.org/10.1016/j.energy.2023.127335.

[173] Zhang H, Tian Y, Tian C, Zhai Z. Effect of key structure and working condition parameters on a compact flat-evaporator loop heat pipe for chip cooling of data centers. Energy 2023;284:128658. https://doi.org/10.1016/j.energy.2023.128658.

[174] He Z, Xi H, Ding T, Wang J, Li Z. Energy efficiency optimization of an integrated heat pipe cooling system in data center based on genetic algorithm. Appl Therm Eng 2021;182:115800. https://doi.org/10.1016/j.applthermaleng.2020.115800.

[175] Xue ZH, Ai BC, Qu W. Data center energy conservation study utilizing loop heat pipes as a chip-level cooling technique and its industrial application. Appl Therm Eng 2023;220:119715. https://doi.org/10.1016/j.applthermaleng.2022.119715.

[176] Weng C, et al. Numerical and experimental investigations of the micro-channel flat loop heat pipe (MCFLHP) heat recovery system for data centre cooling and heat recovery. J Build Eng 2020;35:102088. https://doi.org/10.1016/j.jobe.2020.102088.

[177] 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 Prog 2023;45:102115. https://doi.org/10.1016/j.tsep.2023.102115.

[178] Tang Y, Zhang X, Liu Z. Experimental study on the thermal performance of flat loop heat pipe applied in data center cooling. Energies 2023;16(12). https://doi.org/10.3390/en16124677.

[179] Xiang J, et al. Experimental investigation of a novel micro-channel flat loop heat pipe (MC-FLHP) for data center cooling and heat recovery. Int J Green Energy 2022;19(12):1318–32. https://doi.org/10.1080/15435075.2021.1994978.

[180] Li C, Li J. Thermal characteristics of a flat plate pulsating heat pipe module for onsite cooling of high power server CPUs. Therm Sci Eng Prog 2023;37. https://doi.org/10.1016/j.tsep.2022.101542.

[181] Iceotope, An introduction to how data centers are cooled, 2024. https://www.iceotope.com/learning-hub/insights/an-introduction-to-how-data-centers-are-cooled/ (accessed Mar. 23, 2025).

[182] Wang N, Guo Y, Huang C, Shao S. Advances in direct liquid cooling technology and waste heat recovery for data center: a state-of-the-art review. J Clean Prod 2024;477:143872. https://doi.org/10.1016/j.jclepro.2024.143872.

[183] Ling YZ, Zhang XS, Wang F, She XH. Performance study of phase change materials coupled with three-dimensional oscillating heat pipes with different structures for electronic cooling. Renew Energy 2020;154:636–49. https://doi.org/10.1016/j.renene.2020.03.008.

[184] Chen H, Han Y, Tang G, Zhang X. A dynamic control system for server processor direct liquid cooling. IEEE Trans Components, Packag Manuf Technol 2020;10(5):786–94. https://doi.org/10.1109/TCPMT.2020.2986796.

[185] Anisha, Kumar A. Identification and mitigation of shortcomings in direct and indirect liquid cooling-based battery thermal management system. Energies 2023;16(9). https://doi.org/10.3390/en16093857.

[186] Liu J, Chen H, Huang S, Jiao Y, Chen M. Recent progress and prospects in liquid cooling thermal management system for lithium-ion batteries. Batteries 2023;9(8). https://doi.org/10.3390/batteries9080400.

[187] Shahi P, Mathew A, Akash A, Saini S, Kasukurthy R, Agonafer D. Assessment of reliability enhancement in high-power CPUs and GPUs using dynamic direct-to-chip liquid cooling. J Enhanc Heat Transf 2022;29(8):1–13. https://doi.org/10.1615/JEnhHeatTransf.2022043858.

[188] Kong R, Zhang H, Tang M, Zou H, Tian C, Ding T. Enhancing data center cooling efficiency and ability: a comprehensive review of direct liquid cooling technologies. Energy 2024;308:132846. https://doi.org/10.1016/j.energy.2024.132846.

[189] Chainer TJ, Schultz MD, Parida PR, Gaynes MA. Improving data center energy efficiency with advanced thermal management. IEEE Trans Components, Packag Manuf Technol 2017;7(8):1228–39. https://doi.org/10.1109/TCPMT.2017.2661700.

[190] F. Rebarber, Quantifying the Impact on PUE and Energy Consumption When Introducing Liquid Cooling Into an Air-cooled Data Center, VERTIV, 2023. https://www.vertiv.com/en-asia/about/news-and-insights/articles/blog-posts/quantifying-data-center-pue-when-introducing-liquid-cooling/ (accessed Jan. 23, 2025).

[191] Naduvilakath-Mohammed FM, Jenkins R, Byrne G, Robinson AJ. Closed loop liquid cooling of high-powered CPUs: a case study on cooling performance and energy optimization. Case Stud Therm Eng 2023;50:103472. https://doi.org/10.1016/j.csite.2023.103472.

[192] Hewlett Packard Enterprise (HPE), "The industry's first 100% fanless direct liquid cooling system architecture: Future-focused cooling for optimizing compute-intensive AI and HPC workloads," London, United Kingdom, 2024.

[193] Robison C. Direct liquid cooling on el capitan, the world's most powerful supercomputer. Coolit Systems 2024 (accessed Jan. 23, 2025).

[194] Garimella SV, Persoons T, Weibel JA, Gektin V. Electronics thermal management in information and communications technologies: challenges and future directions. IEEE Trans Components, Packag Manuf Technol 2017;7(8):1191–205. https://doi.org/10.1109/TCPMT.2016.2603600.

[195] Varma D. Data center cold wars – part 3: single-phase immersion cooling versus cold plate. Green Revolution Cooling 2020. https://www.grcooling.com/blog/data-center-cold-wars-part-3-single-phase-immersion-cooling-versus-cold-plate/ (accessed Feb. 03, 2025).

[196] Sunbird Software, Cold Plate Cooling, 2025. https://www.sunbirddcim.com/glossary/cold-plate-cooling (accessed Feb. 03, 2025).

[197] Trumonytechs, Overcoming Challenges in Data Centers With Innovative Cold Plate Cooling Solutions, 2022. https://www.trumonytechs.com/overcoming-challenges-in-data-centers/ (accessed Feb. 03, 2025).

[198] Berktold M, et al. Leak detection and intervention white paper. USA: Ohio; 2021.

[199] Jones Lang LaSalle IP, Getting data centers ready for the AI boom: With the rapid improvements in AI capabilities, are data centers doing enough to keep pace?, 2024. https://www.jll.co.uk/en/trends-and-insights/workplace/getting-data-centers-ready-for-the-ai-boom (accessed Feb. 03, 2025).

[200] Fan S, Duan F. A review of two-phase submerged boiling in thermal management of electronic cooling. Int J Heat Mass Transf 2020;150:119324. https://doi.org/10.1016/j.ijheatmasstransfer.2020.119324.

[201] Kandasamy R, Ho JY, Liu P, Wong TN, Toh KC, Chua SJ. Two-phase spray cooling for high ambient temperature data centers: evaluation of system performance. Appl Energy 2022;305:117816. https://doi.org/10.1016/j.apenergy.2021.117816.

[202] Yang H, Rong L, Liu X, Liu L, Fan M, Pei N. Experimental research on spray evaporative cooling system applied to air-cooled chiller condenser. Energy Rep 2020;6:906–13. https://doi.org/10.1016/j.egyr.2020.04.001.

[203] Yang P, Ling W, Tian K, Zeng M, Wang Q. Flow distribution and heat transfer performance of two-phase flow in parallel flow heat exchange system. Energy 2023;270:126957. https://doi.org/10.1016/j.energy.2023.126957.

[204] Meng X, Meng L, Gao Y, Li H. A comprehensive review on the spray cooling system employed to improve the summer thermal environment: application efficiency, impact factors, and performance improvement. Build Environ 2022;217:109065. https://doi.org/10.1016/j.buildenv.2022.109065.

[205] Wang YF, Wu JT. Thermal performance predictions for an HFE-7000 direct flow boiling cooled battery thermal management system for electric vehicles. Energy Convers Manag 2020;207:112569. https://doi.org/10.1016/j.enconman.2020.112569.

[206] Yin J, et al. Spray cooling as a high-efficient thermal management solution: a review. Energies 2022;15(22):1–29. https://doi.org/10.3390/en15228547.

[207] Asim M, Siddiqui FR. Hybrid nanofluids—next-generation fluids for spray-cooling-based thermal management of high-heat-flux devices. Nanomaterials 2022;12(3):1–29. https://doi.org/10.3390/nano12030507.

[208] Chen H, Ruan XH, Peng YH, Wang YL, Yu CK. Application status and prospect of spray cooling in electronics and energy conversion industries. Sustain Energy

Technol Assessments 2022;52:102181. https://doi.org/10.1016/j.seta.2022.102181.

[209] Bao J, Wang Y, Xu X, Niu X, Liu J, Qiu L. Analysis on the influences of atomization characteristics on heat transfer characteristics of spray cooling. Sustain Cities Soc 2019;51:101799. https://doi.org/10.1016/j.scs.2019.101799.

[210] Liu P, Kandasamy R, Ho JY, Wong TN, Toh KC. Dynamic performance analysis and thermal modelling of a novel two-phase spray cooled rack system for data center cooling. Energy 2023;269:126835. https://doi.org/10.1016/j.energy.2023.126835.

[211] Wu WL, Cheng CG, Li Y, Wei SF, Chen DL. Effect of droplet characteristics on liquid-phase distribution in spray zone of internal mixing air-mist nozzle. China Foundry 2024;21(2):185–96. https://doi.org/10.1007/s41230-024-3059-0.

[212] Mao R, et al. Experimental investigation on the application of cold-mist direct evaporative cooling in data centers. Int J Therm Sci 2025;208:109500. https://doi.org/10.1016/j.ijthermalsci.2024.109500.

[213] Gao N, Ali SM, Persoons T. On the numerical investigation of two-phase evaporative spray cooling technology for data centre applications. Fluids 2024;9(12). https://doi.org/10.3390/fluids9120284.

[214] Cao G, et al. Experimental study on the effect of tube surface modifications and in-channel baffle plates on falling film evaporator heat transfer properties for sever cabinets. Int Commun Heat Mass Transf 2024;158:107943. https://doi.org/10.1016/j.icheatmasstransfer.2024.107943.

[215] Klinkhamer C, Abishek S, Iyer KLV, Balachandar R, Barron R. Characterization of a jet impingement heat sink for power electronics cooling. Therm Sci Eng Prog 2022;34:101408. https://doi.org/10.1016/j.tsep.2022.101408.

[216] Cai Z, Wang C, Luo Q, Chen W. Enhancing heat transfer characteristics and optimization of a multi-jet cooling system. Appl Therm Eng 2024;248:123288. https://doi.org/10.1016/j.applthermaleng.2024.123288.

[217] Wu Z, et al. Jet microchannel with sawtooth wall for efficient cooling of high-power electronics. Int J Heat Mass Transf 2023;206. https://doi.org/10.1016/j.ijheatmasstransfer.2023.123955.

[218] Joshi SN, et al. A review of select patented technologies for cooling of high heat flux power semiconductor devices. IEEE Trans Power Electron 2023;38(6):6790–4. https://doi.org/10.1109/TPEL.2023.3243546.

[219] Azarifar M, Arik M. Liquid synthetic jets for high flux electronics cooling. Appl Therm Eng 2025;261:125007. https://doi.org/10.1016/j.applthermaleng.2024.125007.

[220] Sarkar S, Gupta R, Roy T, Ganguly R, Megaridis CM. Review of jet impingement cooling of electronic devices: emerging role of surface engineering. Int J Heat Mass Transf 2023;206. https://doi.org/10.1016/j.ijheatmasstransfer.2023.123888.

[221] Lu Q, Muthukumar R, Ge H, Parameswaran S. Numerical study of a rotating liquid jet impingement cooling system. Int J Heat Mass Transf 2020;163. https://doi.org/10.1016/j.ijheatmasstransfer.2020.120446.

[222] Gil P, Wilk J, Smusz R, Gałek R. Centerline heat transfer coefficient distributions of synthetic jets impingement cooling. Int J Heat Mass Transf 2020;160. https://doi.org/10.1016/j.ijheatmasstransfer.2020.120147.

[223] Jung J-Y, Choi W-W, Kim S-M. Jet array impingement heat transfer in a rectangular cavity with effusion holes. Int J Mech Sci 2024;283:109698. https://doi.org/10.1016/j.ijmecsci.2024.109698.

[224] Zhang Y, Wang Y, Kosonen R, Khan DA, Wang X. Numerical research on the flow and heat transfer characteristics in the immersion jet cooling for servers. Case Stud Therm Eng 2024;60:104748. https://doi.org/10.1016/j.csite.2024.104748.

[225] Hefny M, Cotton JS, Emadi A. The effects of outlet size, shape, and location on spatial temperature distribution reduction via direct jet impinging in power electronics. Appl Therm Eng 2024;249:123439. https://doi.org/10.1016/j.applthermaleng.2024.123439.

[226] Kaood A, Elhagali IO, Hassan MA. Investigation of high-efficiency compact jet impingement cooling modules for high-power applications. Int J Therm Sci 2023;184:108006. https://doi.org/10.1016/j.ijthermalsci.2022.108006.

[227] Nemati K, Alissa HA, Murray BT, Schneebeli K, Sammakia B. Experimental failure analysis of a rear door heat exchanger with localized containment. IEEE Trans Components, Packag Manuf Technol 2017;7(6):882–92. https://doi.org/10.1109/TCPMT.2017.2682863.

[228] Michaelis A. The flexibility of rear door heat exchangers in data centers. STACK Infrastruct 2024. https://www.stackinfra.com/resources/blog/the-flexibility-of-rear-door-heat-exchangers-in-data-centers/ (accessed Mar. 23, 2025).

[229] Burse J. Rear door heat exchangers (RDHx). Prasa Infocom & Power Solutions Pvt Ltd. 2023 https://prasa-pl.com/blog/rear-door-heat-exchangers-rdhx/ (accessed Mar. 28, 2025).

[230] Cho J. Numerical analysis of rack-based data center cooling with rear door heat exchanger (RDHx): Interrelationship between thermal performance and energy efficiency. Case Stud Therm Eng 2024;63:105247. https://doi.org/10.1016/j.csite.2024.105247.

[231] Yang W, Yang L, Ou J, Lin Z, Zhao X. Investigation of heat management in high thermal density communication cabinet by a rear door liquid cooling system. Energies 2019;12(22). https://doi.org/10.3390/en12224385.

[232] Zhou F, Li C, Zhu W, Zhou J, Ma G, Liu Z. Energy-saving analysis of a case data center with a pump-driven loop heat pipe system in different climate regions in China. Energ Build 2018;169:295–304. https://doi.org/10.1016/j.enbuild.2018.03.081.

[233] Ding T, Guang He Z, Hao T, Li Z. Application of separated heat pipe system in data center cooling. Appl Therm Eng 2016;109:207–16. https://doi.org/10.1016/j.applthermaleng.2016.08.025.

[234] Hnayno M, Chehade A, Klaba H, Bauduin H, Polidori G, Maalouf C. Performance analysis of new liquid cooling topology and its impact on data centres. Appl Therm Eng 2022;213:118733. https://doi.org/10.1016/j.applthermaleng.2022.118733.

[235] Han Z, Xue D, Wei H, Ji Q, Sun X, Li X. Study on operation strategy of evaporative cooling composite air conditioning system in data center. Renew Energy 2021;177:1147–60. https://doi.org/10.1016/j.renene.2021.06.046.

[236] Wang P, et al. District heating utilizing waste heat of a data center: high-temperature heat pumps☆. Energ Buildings 2024;315:114327. https://doi.org/10.1016/j.enbuild.2024.114327.

[237] Gao T, Murray B, Sammakia B. Analysis of transient and hysteresis behavior of cross-flow heat exchangers under variable fluid mass flow rate for data center cooling applications. Appl Therm Eng 2015;84:15–26. https://doi.org/10.1016/j.applthermaleng.2015.03.044.

[238] Li X, Li M, Zhang Y, Han Z, Wang S. Rack-level cooling technologies for data centers – a comprehensive review. J Build Eng 2024;90:109535. https://doi.org/10.1016/j.jobe.2024.109535.

[239] H. Ebermann, Immersion cooling systems: Advantages and deployment strategies for AI and HPC data centers, 2024. https://www.vertiv.com/en-emea/about/news-and-insights/articles/blog-posts/advancing-data-center-performance-with-immersion-cooling/ (accessed Apr. 12, 2025).

[240] A. von Hassler, CRAC vs. CRAH Cooling Units: What's the Difference?, DataSpan, 2021. https://dataspan.com/blog/crac-vs-crah-cooling-units-whats-the-difference/ (accessed Aug. 21, 2024).

[241] Hyve Limited, Understanding Data Centre Cooling: Methods and Solutions Used in Our Data Centres. https://www.hyve.com/insights/how-are-data-centres-cooled/ (accessed Mar. 18, 2025).

[242] Microsoft, Modern datacenter cooling, 2023. [Online]. Available: https://datacenters.microsoft.com/wp-content/uploads/2023/05/Azure_Modern-Datacenter-Cooling_Infographic.pdf.

[243] Law M. Microsoft unveils zero-water cooling for AI data centres. Data Centre Magazine 2024. https://datacentremagazine.com/articles/microsoft-unveils-zero-water-cooling-for-ai-data-centres (accessed Mar. 20, 2025).

[244] Zhang M. Data center water usage: a comprehensive guide. Dgtl Infra 2024. https://dgtlinfra.com/data-center-water-usage/ (accessed Apr. 13, 2025).

[245] ARANER, Data Centers cooling methods: for the Next Online Revolution, 2025. https://www.araner.com/blog/best-data-center-cooling-system (accessed Mar. 13, 2025).

[246] Osias T, Schulz L. Exploring liquid cooling for next-gen business applications. Equinix 2023. https://blog.equinix.com/blog/2023/09/26/exploring-liquid-cooling-for-next-gen-business-applications/ (accessed Mar. 18, 2025).

[247] CoolIT Systems, How Direct Liquid Cooling Improves Data Center Energy Efficiency, Calgary, Canada, 2024.

[248] Walsh N. Sustainable by design: transforming datacenter water efficiency. Microsoft 2024. https://www.microsoft.com/en-us/microsoft-cloud/blog/2024/07/25/sustainable-by-design-transforming-datacenter-water-efficiency/ (accessed Mar. 18, 2025).

[249] O'Gorman M, Metcalf G. Data center cooling continues to evolve for efficiency and density. Equinix 2023. https://blog.equinix.com/blog/2023/12/11/data-center-cooling-continues-to-evolve-for-efficiency-and-density/ (accessed Mar. 18, 2025).

[250] Mann T. More than a third of enterprise datacenters expect to deploy liquid cooling by 2026. Register 2024. https://www.theregister.com/2024/04/22/register_liquid_cooling_survey/ (accessed Mar. 23, 2025).

[251] Submer, Immersion Cooling at Scale: Navigate evolving workload demands effortlessly, 2025. https://submer.com/ (accessed Mar. 23, 2025).

[252] Williams A. Optimizing energy efficiency in datacenters with advanced cooling technologies: managing the growing power requirements of datacenter thermal management. accessed Mar. 18, 2025 Communications 2024. https://cacm.acm.org/blogcacm/optimizing-energy-efficiency-in-datacenters-with-advanced-cooling-technologies/.

[253] Reuters, Schneider Electric to buy data centre cooling firm Motivair for $850 million, 2024. https://www.reuters.com/markets/deals/schneider-electric-buy-data-centre-cooling-firm-motivair-850-million-2024-10-17/ (accessed Mar. 18, 2025).

[254] Data Center Frontier, nVent Collaborates with NVIDIA on AI-Ready Liquid Cooling Solutions, 2024. https://www.datacenterfrontier.com/press-releases/article/55245330/nvent-collaborates-with-nvidia-on-ai-ready-liquid-cooling-solutions (accessed Mar. 18, 2025).

[255] Chernicoff D, Vincent M. Supermicro rises to meet massive growth in data center liquid cooling demand. Data Center Front 2024. https://www.datacenterfrontier.com/machine-learning/article/55128681/supermicro-rises-to-meet-massive-growth-in-data-center-liquid-cooling-demand (accessed Mar. 18, 2025).

[256] WTWH Media, HPE Launches New Purpose-Built Solutions Powered by AMD, 2024. https://www.engineering.com/hpe-launches-new-purpose-built-solutions-powered-by-amd/ (accessed Mar. 18, 2025).

[257] Chernicoff D, Vincent M. Global data center operator telehouse launches liquid cooling lab in the UK to meet ongoing AI and HPC demand. Data Center Front 2025. https://www.datacenterfrontier.com/cooling/article/55246042/vertiv-prepares-for-future-of-air-and-liquid-cooling-taps-compass-datacenters-as-key-partner (accessed Mar. 18, 2025).

[258] Boyd, Enhancing AI: Deploy NVIDIA's GB200 NVL72 Superchip Easier and Faster with Boyd's Modular Liquid Cooling Systems, 2024. https://www.boydcorp.com/blog/cooling-nvidia-gb200-nvl72-artificial-intelligence.html (accessed Mar. 23, 2025).

[259] Swinhoe D. Lenovo reveals liquid-cooled Nvidia Blackwell server. Data Centre Dynamics Ltd (DCD) 2024. https://www.datacenterdynamics.com/en/news/lenovo-reveals-liquid-cooled-nvidia-blackwell-server/ (accessed Mar. 23, 2025).

[260] Shilov A. Intel and submer develop immersion cooling for 1000W CPUs. Tom’s Hardware 2023. https://www.tomshardware.com/news/intel-and-submer-develop-immersion-cooling-for-1000w-cpus (accessed Mar. 23, 2025).

[261] Submer, Immersion Cooling at Scale: Navigate evolving workload demands effortlessly, 2025.

[262] Shell, Shell Liquid Cooling Fluids: Improving efficiency in high-performance computing, 2019. https://www.shell.com/business-customers/lubricants-for-business/process-oils/immersion-cooling-fluids.html#iframe=L2wvMjcwMTEyLzIwMTktMTEtMDcvYmNnOTg/ZGg9MTEwMC0yMDAw (accessed Mar. 23, 2025).

[263] N. Singer, Electrical circuits encased in fluid may reshape data-center design, New Mexico, United States, 2024.

[264] Data Centre Dynamics Ltd (DCD), PMC Technologies proudly unveils its latest innovation in data center cooling technology: oil spray cooling, 2024. https://www.datacenterdynamics.com/en/product-news/pmc-technologies-proudly-unveils-its-latest-innovation-in-data-center-cooling-technology-oil-spray-cooling/ (accessed Mar. 27, 2025).

[265] Cooling Post, Spray cooling could reduce data centre cooling costs, 2023. https://www.coolingpost.com/features/spray-cooling-could-reduce-data-centre-cooling-costs/ (accessed Mar. 27, 2025).

[266] Danfoss, A strategic collaboration between Danfoss and HyprCool to accelerate the implementation of liquid cooling solutions within data centres, Data Centre World, 2024. https://www.datacentreworldasia.com/exhibitor-press-release/strategic-collaboration-between-danfoss-hyprcool-accelerate-implementation-liquid-cooling-solutions-within-data-centres?utm_source=chatgpt.com (accessed Mar. 28, 2025).

[267] Flex, Integrated, scalable, and liquid-cooled compute racks for AI data centers, Texas, United States, 2024. Accessed: Mar. 28, 2025. [Online]. Available: https://flex.com/resources/advanced-liquid-cooling-for-ai-data-centers.

[268] EDP Europe, ColdLogik Water Cooled Server Rack Cooling, 2025. https://www.edpeurope.com/products/thermal-management/water-in-row-cooling/coldlogik-water-cooled-server-rack-2/ (accessed Mar. 23, 2025).

[269] EDP Europe, Guide to Water/Liquid Cooling in Data Centres, 2025. https://www.edpeurope.com/data-centre-guide/guide-to-water-liquid-cooling-in-data-centres/ (accessed Mar. 23, 2025).

[270] Hewlett Packard Enterprise (HPE), Motivair Liquid Cooled Doors, Texas, United States, 2022.

[271] Chernicoff D, Vincent M. Vertiv prepares for future of air and liquid cooling, taps compass datacenters as key partner. Data Center Front 2024. https://www.datacenterfrontier.com/cooling/article/55246042/vertiv-prepares-for-future-of-air-and-liquid-cooling-taps-compass-datacenters-as-key-partner (accessed Mar. 18, 2025).

[272] Vertiv, Liquid Cooling Options for Data Centers, 2024. https://www.vertiv.com/en-emea/solutions/learn-about/liquid-cooling-options-for-data-centers/ (accessed Mar. 19, 2025).

[273] Alfa Laval, Heat recovery, 2025. https://www.alfalaval.co.uk/industries/hvac/data-center-cooling/heat-recovery/ (accessed Apr. 13, 2025).

[274] Mytton D. Data centre water consumption. npj Clean Water 2021;4(1). https://doi.org/10.1038/s41545-021-00101-w.

[275] Research and Markets, U.S. Data Center Cooling Market 2024-2029: Rising Adoption of AI to Drive Demand for Advanced Cooling Technologies, 2024. https://www.globenewswire.com/news-release/2024/08/30/2938615/28124/en/U-S-Data-Center-Cooling-Market-2024-2029-Rising-Adoption-of-AI-to-Drive-Demand-for-Advanced-Cooling-Technologies.html (accessed Mar. 11, 2025).

[276] Shenzhen Envicool Technology, XFLEX Indirect Free Cooling Solution, 2025.

[277] Vertiv, Calculating the Impact of Water Usage on Data Center Costs and Sustainability, Ohio, USA, 2020.

[278] S. Krishnapura, J. Musilli, and P. Budhai, Intel IT: extremely energy-efficient, high-density data centers, 2015.

[279] EDP Europe, Guide to Water/Liquid Cooling in Data Centres, 2025.

[280] A. Chehade, M. Hnayno, and J. Jay, Understanding OVHcloud’s data centres: how Delta T helps with water efficiency, 2023. https://blog.ovhcloud.com/understanding-ovhclouds-data-centres-how-delta-t-helps-with-water-efficiency/ (accessed Apr. 13, 2025).

[281] Asgari S, et al. Hybrid surrogate model for online temperature and pressure predictions in data centers. Futur Gener Comput Syst 2021;114:531–47. https://doi.org/10.1016/j.future.2020.08.029.

[282] Zou S, Zhang Q, Yue C. Simulation on the performance and cooling capacity compensation solution of a loop thermosyphon system under fan failure conditions. Case Stud Therm Eng 2024;57:104360. https://doi.org/10.1016/j.csite.2024.104360.

[283] Moazamigoodarzi H, Gupta R, Pal S, Tsai PJ, Ghosh S, Puri IK. Modeling temperature distribution and power consumption in IT server enclosures with row-based cooling architectures. Appl Energy 2020;261:114355. https://doi.org/10.1016/j.apenergy.2019.114355.

[284] Wang N, Guo Y, Huang C, Tian B, Shao S. Multi-scale collaborative modeling and deep learning-based thermal prediction for air-cooled data centers: an innovative insight for thermal management. Appl Energy 2025;377:124568. https://doi.org/10.1016/j.apenergy.2024.124568.

[285] Li G, et al. An explainable one-dimensional convolutional neural networks based fault diagnosis method for building heating, ventilation and air conditioning systems. Build Environ 2021;203:108057. https://doi.org/10.1016/j.buildenv.2021.108057.

[286] Moazamigoodarzi H, Pal S, Ghosh S, Puri IK. Real-time temperature predictions in IT server enclosures. Int J Heat Mass Transf 2018;127:890–900. https://doi.org/10.1016/j.ijheatmasstransfer.2018.08.091.

[287] Rehman R, Malik SUR, Qureshi SH, Mouqarrab SA. A novel multi-modal Federated Learning based thermal-aware job scheduling framework. Comput Netw 2024;255:110879. https://doi.org/10.1016/j.comnet.2024.110879.

[288] Arroyo J, Manna C, Spiessens F, Helsen L. Reinforced model predictive control (RL-MPC) for building energy management. Appl Energy 2022;309:118346. https://doi.org/10.1016/j.apenergy.2021.118346.

[289] Deng W, Wang J, Yue C, Guo Y, Zhang Q. Model-based control strategy with linear parameter-varying state-space model for rack-based cooling data centers. Energ Buildings 2024;319:114528. https://doi.org/10.1016/j.enbuild.2024.114528.

[290] Park BR, Choi YJ, Choi EJ, Moon JW. Adaptive control algorithm with a retraining technique to predict the optimal amount of chilled water in a data center cooling system. J Build Eng 2022;50:104167. https://doi.org/10.1016/j.jobe.2022.104167.

[291] Liu W, Tong X, Wang J, Yue C, Zhang Q. Real-time temperature predictions via state-space model and parameters identification within rack-based cooling data centers. J Build Eng 2022;58:105013. https://doi.org/10.1016/j.jobe.2022.105013.

[292] Asgari S, MirhoseiniNejad SM, Moazamigoodarzi H, Gupta R, Zheng R, Puri IK. A gray-box model for real-time transient temperature predictions in data centers. Appl Therm Eng 2021;185:116319. https://doi.org/10.1016/j.applthermaleng.2020.116319.

[293] Hosseini P, Taheri S, Akhavan J, Razban A. Privacy-preserving federated learning: application to behind-the-meter solar photovoltaic generation forecasting. Energy Convers Manag 2023;283. https://doi.org/10.1016/j.enconman.2023.116900.

[294] Huang Y, Song Y, Jing Z. Load forecasting using federated learning with considering electricity data privacy preservation of EASP. Ain Shams Eng J 2024;15(6):102724. https://doi.org/10.1016/j.asej.2024.102724.

[295] Ali M, Singh AK, Kumar A, Ali SS, Choi BJ. Comparative analysis of data-driven algorithms for building energy planning via federated learning. Energies 2023;16(18). https://doi.org/10.3390/en16186517.

[296] Deepak G, Parthiban M, Nath SS, Sulaiman Alfurhood B, Mouleswararao B, Ravi Kishore V. Ai-enhanced thermal modeling for integrated process-product-system optimization in zero-defect manufacturing chains. Therm Sci Eng Prog 2024;55:102945. https://doi.org/10.1016/j.tsep.2024.102945.

[297] Ranawaka A, Alahakoon D, Sun Y, Hewapathirana K. Leveraging the synergy of digital twins and artificial intelligence for sustainable power grids: a scoping review. Energies 2024;17(21). https://doi.org/10.3390/en17215342.

[298] Ghadim HB, Godin A, Veillere A, Duquesne M, Haillot D. Review of thermal management of electronics and phase change materials. Renew Sustain Energy Rev 2025;208:115039. https://doi.org/10.1016/j.rser.2024.115039.

[299] Furszyfer Del Rio DD, Al Kez D, Hasan Wong FWBM, Dolfi A, Srinivasan G, Foley AM. ‘Encapsulating’ experts’ knowledge: an exploration of benefits, risks, barriers and future opportunities of PCMs. Sustain Energy Technol Assessments 2024;71:103980. https://doi.org/10.1016/j.seta.2024.103980.

[300] Martin MJ, et al. Energy use in quantum data centers: scaling the impact of computer architecture, qubit performance, size, and thermal parameters. IEEE Trans Sustain Comput 2022;7(4):864–74. https://doi.org/10.1109/TSUSC.2022.3190242.

[301] Ajagekar A, You F. Variational quantum circuit learning-enabled robust optimization for AI data center energy control and decarbonization. Adv Appl Energy 2024;14:100179. https://doi.org/10.1016/j.adapen.2024.100179.

[302] Zhao Z, Fan L, Han Z. Optimal Data center energy management with hybrid quantum-classical multi-cuts benders’ decomposition method. IEEE Trans Sustain Energy 2024;15(2):847–58. https://doi.org/10.1109/TSTE.2023.3309244.

[303] Fattah TR, Jalal SJ, Katona TJ, Baranyai B. Bio-curvilinear shading device for cooling load reduction in office buildings assessed by a novel energy simulation framework. Results Eng 2024;23. https://doi.org/10.1016/j.rineng.2024.102493.

[304] Usman MJ, et al. Energy-efficient nature-inspired techniques in cloud computing datacenters. Telecommun Syst 2019;71(2):275–302. https://doi.org/10.1007/s11235-019-00549-9.

[305] A. O’Driscoll, Peltier vs. Compressor-Based Cooling, Incubators, 2019. https://labincubators.net/blogs/blog/peltier-vs-compressor-based-cooling (accessed Apr. 12, 2025).

[306] Bu Z, et al. A record thermoelectric efficiency in tellurium-free modules for low-grade waste heat recovery. Nat Commun 2022;13(1):1–8. https://doi.org/10.1038/s41467-021-27916-y.